Télécom Physique Strasbourg

Master IRIV

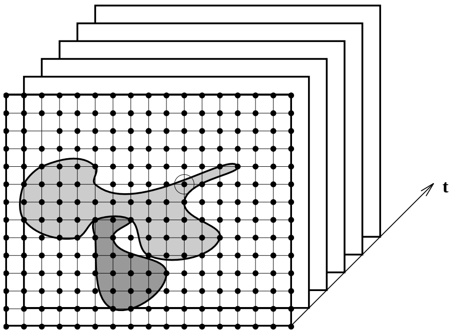


IMAGE SEQUENCE ANALYSIS
M 36032

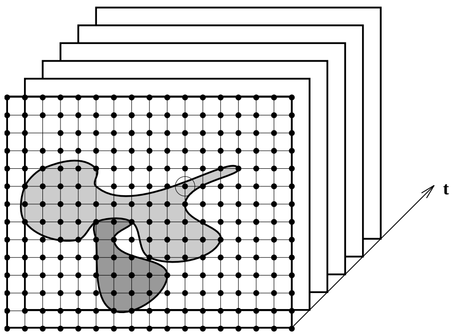
Pr. Fabrice HEITZ

Télécom PS / ICube CNRS

January 2015

Introduction

IMAGE SEQUENCE ANALYSIS



Lecture Outline

Image Sequence Analysis

1. 2D motion detection
2. 2D motion estimation (measurement)
3. 3D structure and motion from 2D motion
4. 2D/3D motion tracking

Detailed plan I

1. **2D motion detection**
    1. Temporal change detection : image difference, hypothesis test
    2. Reconstruction of moving objects masks
    3. Reference image-based techniques
2. **2D motion estimation (measurement)**
    1. The aperture problem, motion field, optical flow
    2. Motion by correspondence : matching of points of interest, SIFT, edges, regions, graphs
    3. Differential approaches : definition of the Displaced Frame Difference (DFD), direct minimization of the DFD (pel recursive methods, block matching), gradient-based methods, parametric motion models, quadratic and semi-quadratic regularization

Detailed plan II

3. **3D structure and motion from 2D motion**
   1. Elements of visual perception
   2. Pinhole camera model
   3. 2D-3D structure and motion equations : equations in position and velocity
   4. Applications
4. **2D/3D motion tracking**
   1. Short term tracking : correspondence
   2. Long term tracking 1 : The Kalman filter
   3. Case study : tracking 2D image features
   4. Long term tracking 2 : The particle filter (sequential Monte Carlo method)
   5. Applications

Web references

- *CVonline : The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision,* http ://homepages.inf.ed.ac.uk/rbf/CVonline
- *CVonline : Motion and Time Sequence Analysis,* http ://homepages.inf.ed.ac.uk/rbf/CVonline/motion.htm
- *CVonline : Vision Related Books including Online Books and Book Support Sites,* http ://homepages.inf.ed.ac.uk/rbf/CVonline/books.htm

References

S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M. Black, R. Szeliski, *A Database and Evaluation Methodology for Optical Flow*, Int. J. Computer Vision, Vol. 92, 2011, pp. 1-31.

J.L. Barron, D.J. Fleet, S.S. Beauchemin, *Performances of Optical Flow Techniques*, Int. J. Computer Vision, Vol. 12, No 1, 1994, pp. 43-77.

A. Bovik (Editor), *Handbook of Image and Video Processing*, Academic Press, 2000.

R. Deriche and O. Faugeras. *Tracking line segments*. Image and Vision Computing, Vol. 8, No 4, pp. 261-270, 1990.

B.K.P. Horn, *Robot Vision*, MIT Press, 1986.

M. Isard and A. Blake, condensation – *Conditional Density Propagation for Visual Tracking*, Int. J. Computer Vision, Vol. 29, No 1, pp. 5-28, 1998.

V. Lepetit, P. Fua, *Monocular Model-Based 3D Tracking of Rigid Objects : A Survey*, Foundations and Trends in Computer Graphics and Vision, Vol. 1, No 1, pp. 1-89, 2005.

A. Mitiche, P. Bouthemy, *Computation and Analysis of Image Motion : A Synopsis of Current Problems and Methods*, Int. Jl. of Computer Vision, Vol. 19, No 1, 1996, pp. 29-55.

References

V. Noblet, C. Heinrich, F. Heitz, J.-P. Armspach, *Recalage d'images médicales*, Techniques de l'Ingénieur, Dossier med910, 11/2013.

I. Pitas, *Digital video and television*, I. Pitas Editor, 2014.

A. Sotiras, C. Davatzikos, N. Paragios, *Deformable Medical Image Registration*, IEEE Trans. Medical Imaging, Vol. 32, No 7, July 2013.

R. Szeliski, *Image Alignement and Stitching : A Tutorial* , Foundations and Trends in Computer Graphics and Vision, Vol. 2, No 1, pp. 1-104, 2006.

R. Szeliski, *Computer Vision : Algorithms and Applications*, Texts in Computer Sciences Series, Springer, 2011. `http://szeliski.org/Book/`

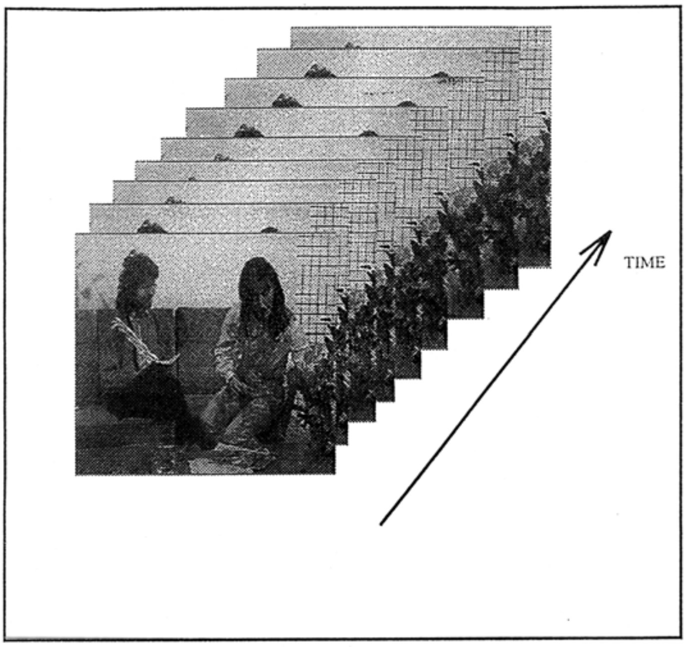
A. M. Tekalp, *Digital Video Processing*, Prentice Hall, 1995.

E. Trucco, A. Verri, *Introductory Techniques for 3D Computer Vision*, Prentice Hall, 1998.

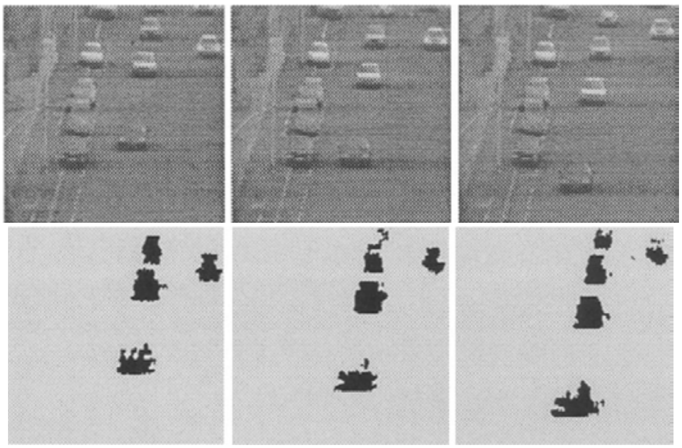
G. Tziritas, C. Labit, *Motion Analysis for Image Sequence Coding*, North-Holland, 1994.

A. Yilmaz, O. Javed and M. Shah, *Object tracking : A survey*, ACM Journal of Computing Surveys, Vol. 38, No 4, 2006.

Image sequence analysis



I - 2D motion analysis : motion detection



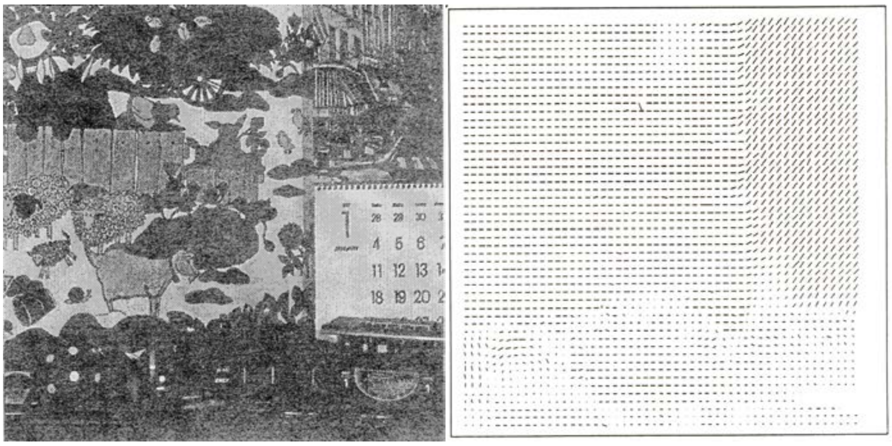
P. Bouthemy, P. Lalande, ECCV 1990

- Binary decision (motion or no motion)
- Goal : reconstruction of moving objects masks
- Approaches :
  - camera motion compensation ;
  - temporal change detection ;
  - background subtraction (change detection / reference image)

T. Veit, F. Cao, P. Bouthemy, An a contrario Decision Framework for Region-Based Motion Detection, IJCV, 2006

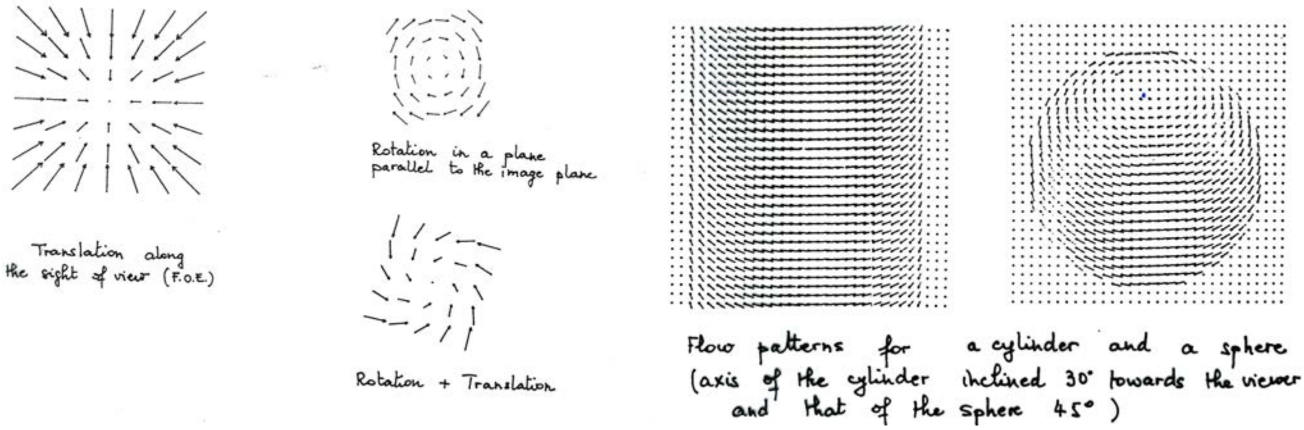
Region-based (level lines) motion detection

II - 2D motion analysis : motion estimation



F. Heitz, P. Bouthemy, Multimodal estimation of discontinuous optical flow using Markov Random Fields, IEEE Trans. PAMI, Vol. 15, No 12, pp. 1217-1232, Dec. 1993

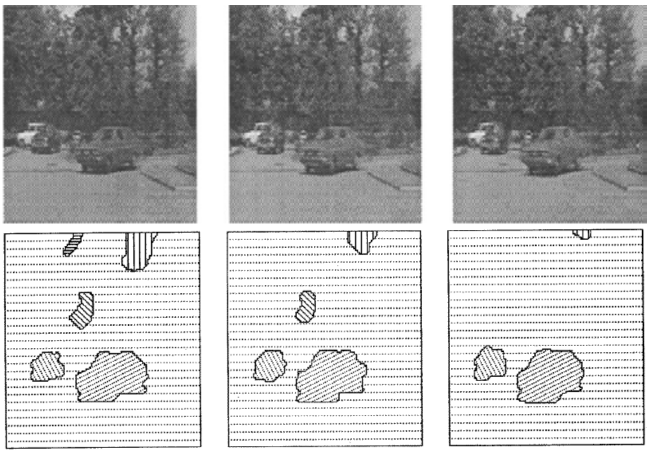
- Goal : apparent motion measurement (=optical flow)
- Approaches :
  - image features matching ;
  - differential approaches ;
  - space-frequency filters.



Some typical exact optical flows

P. Bouthemy, Irisa, 1994

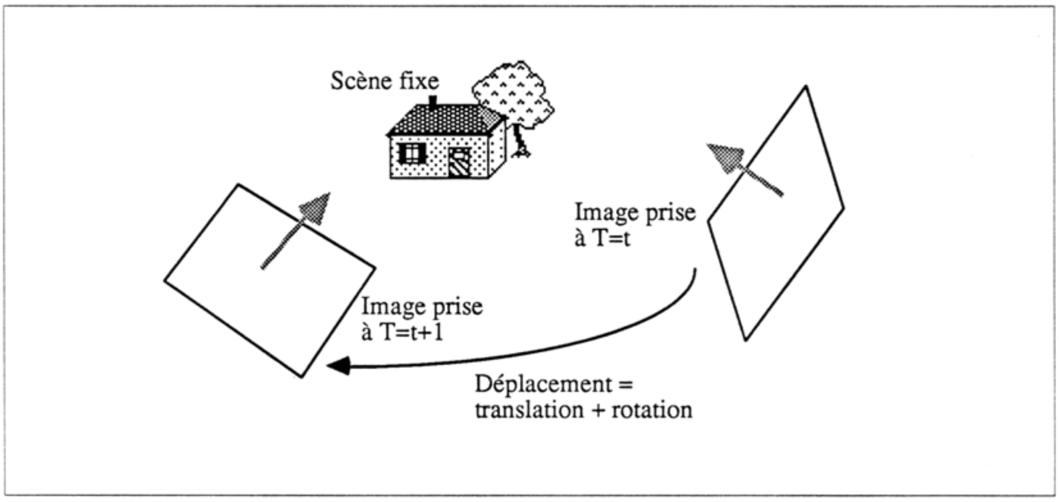
III - 2D motion analysis : motion segmentation



P. Bouthemy, E. Francois. Motion segmentation and qualitative dynamic scene analysis from an image sequence. Int. Journal of Computer Vision, Vol. 10(2), pp. 157-182, April 1993

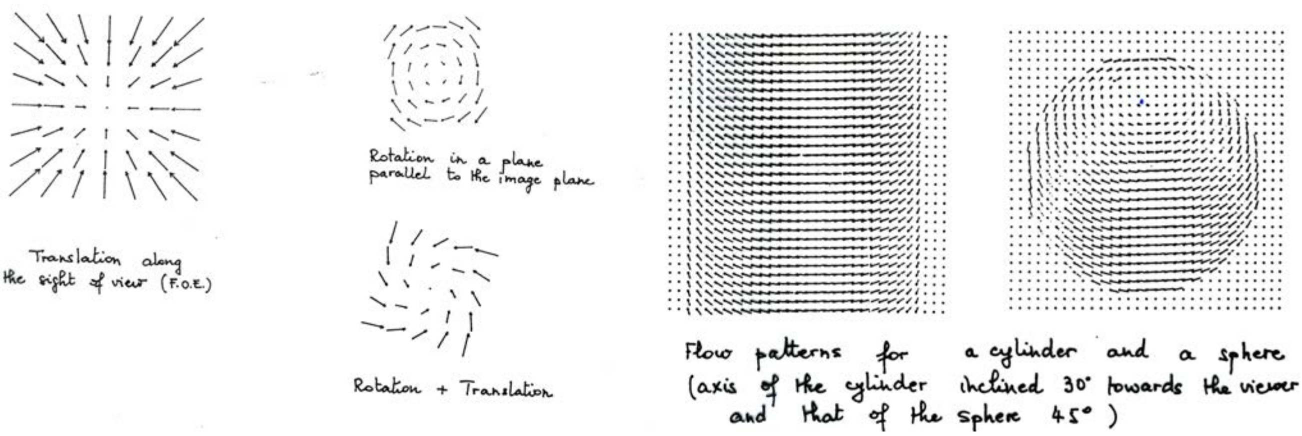
- Goal : segmentation in regions with different motion
- Approaches :
  - segmentation from optical flow field ;
  - segmentation from spatio-temporal gradients ;
  - combination of spatial and motion segmentation.

IV - 3D structure and motion from 2D motion



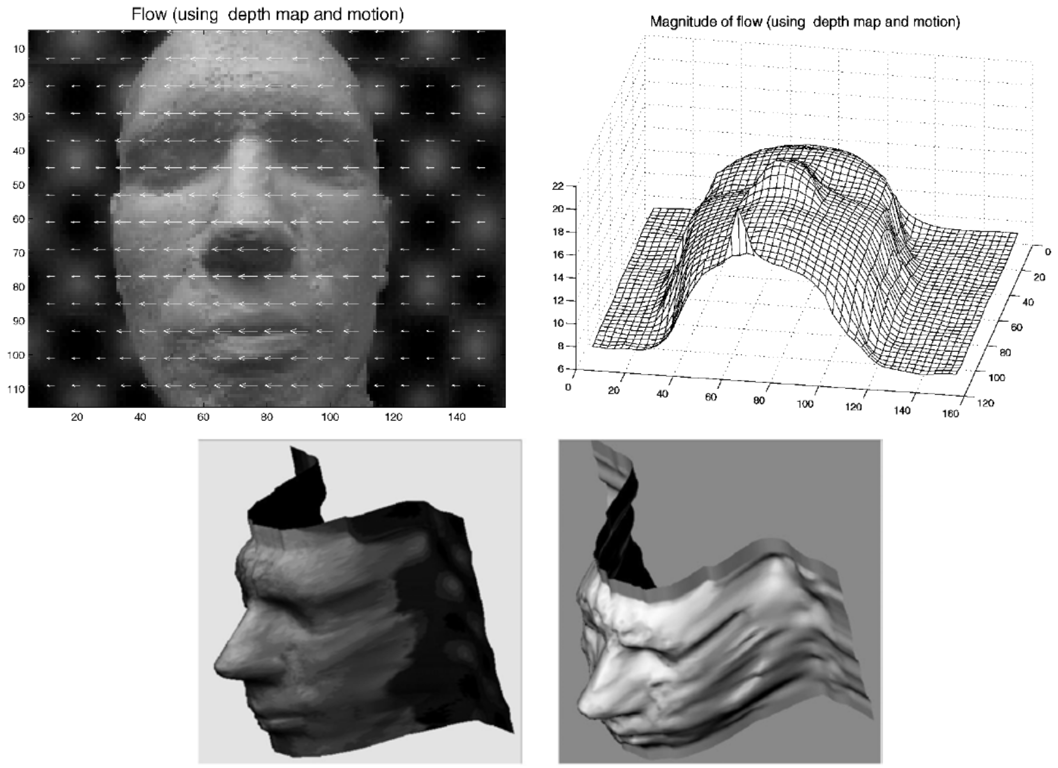
L. Pocztar, Thèse Univ. Paris 6, 1987

- Goal : reconstruction of 3D structure and motion (rigid case) from 2D motion (measured in the image plane).
- Approaches :
  - based on measured 2D motion field (optical flow) ;
  - based on matching of image features.



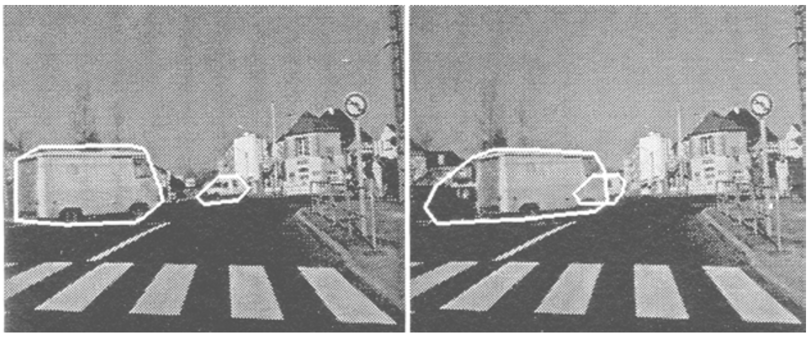
Some typical exact optical flows

P. Bouthemy, Irisa, 1994



G.P. Stein, A. Shashua, Model-Based Brightness Constraints : On Direct Estimation of Structure and Motion, IEEE Trans. PAMI, Sept. 2000

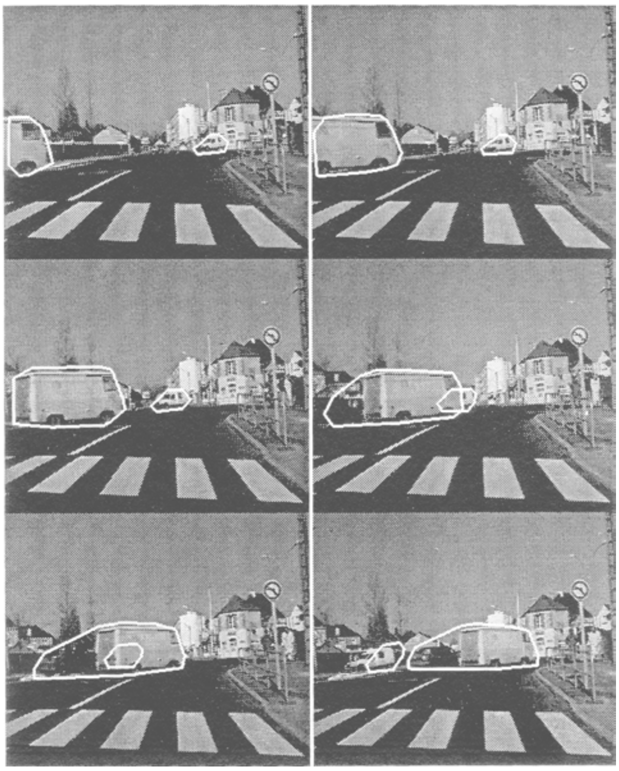
V - 2D/3D motion tracking



F. Meyer, Thèse Univ. Rennes I, 1993

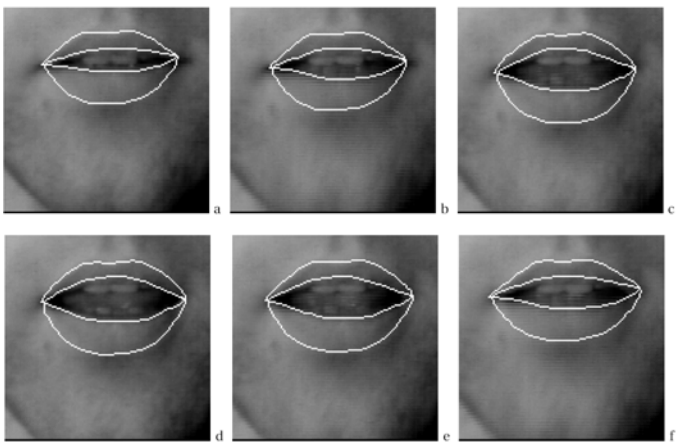
- Goal : tracking of segmented regions over long image sequences, trajectography.
- Approaches :
  - feature matching, block matching ;
  - recursive temporal filtering (Kalman filter and variants) ;
  - particle filters (sequential Monte Carlo methods) [Isard and Blake 1996]

V - 2D/3D motion tracking



F. Meyer, Thèse Univ. Rennes I, 1993

V - 2D/3D motion tracking

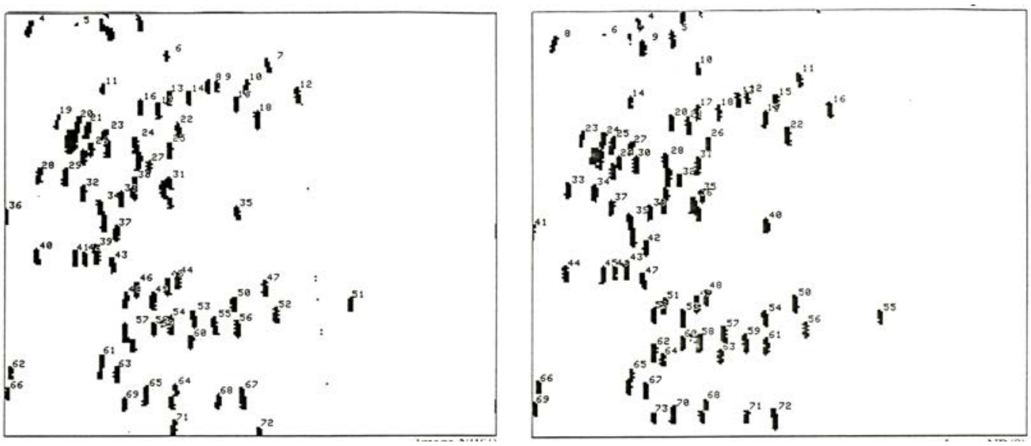


C. Kervrann, F. Heitz, A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes,

CVGIP : Graphical Models and Image Processing, pp. 173-195, Vol. 60, Num. 3, 1998

Tracking of 2D deformable structures

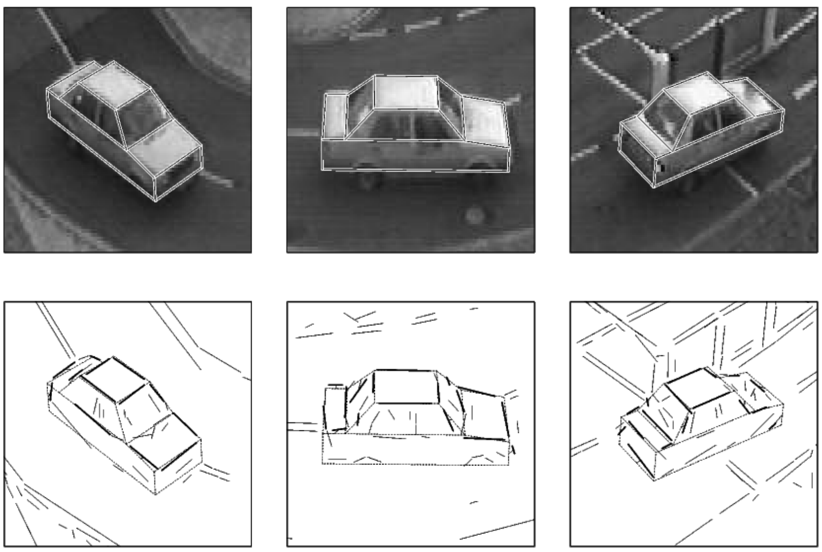
V - 2D/3D motion tracking



L. Pocztar, Thèse Univ. Paris 6, 1987

Tracking of a particle flow

V - 2D/3D motion tracking



D. Koller, K. Daniilidis, and H.-H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, IJCV, vol. 10, pp. 257-281, July 1993.

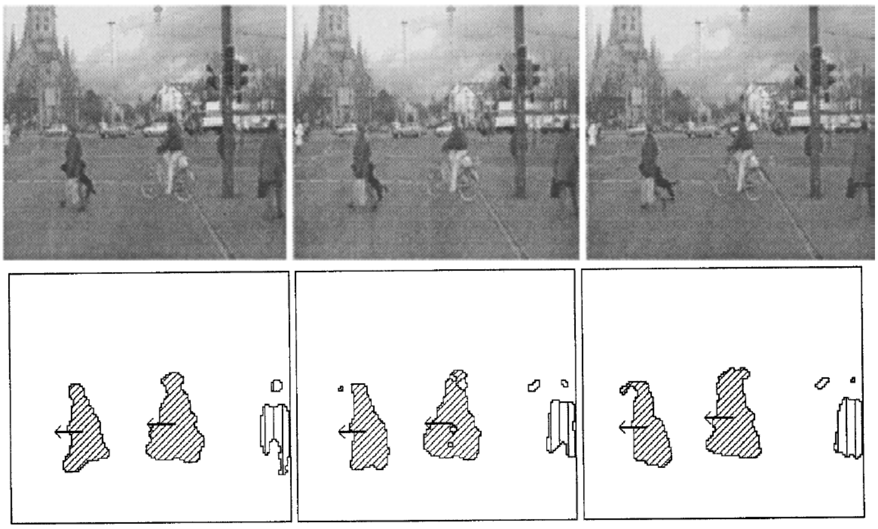
Tracking of a 3D rigid model (car)

V - 2D/3D motion tracking

L. Vacchetti, V. Lepetit, and P. Fua, Stable real-time 3D tracking using online and offline information, IEEE Trans. PAMI, vol. 26, pp. 1385-1391, October 2004.

Tracking of a 3D deformable model (face)

VI - 2D/3D Motion interpretation



P. Bouthemy, E. Francois. Motion segmentation and qualitative dynamic scene analysis from an image sequence. Int. Journal of Computer Vision, Vol. 10(2), pp. 157-182, April 1993

- Goal : interpretation of objects motion
- Approaches :
  - pattern recognition techniques ;
  - quantitative / qualitative interpretation

Télécom Physique Strasbourg Module 36032

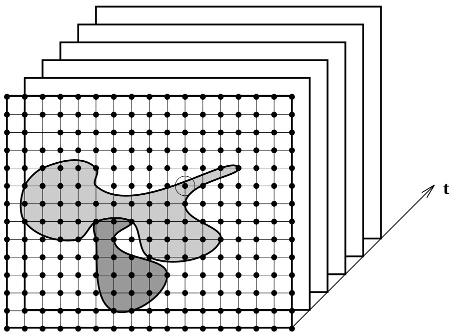
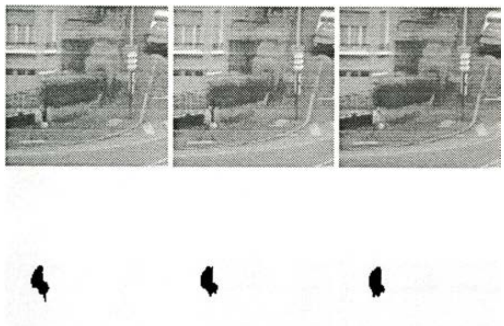


IMAGE SEQUENCE ANALYSIS

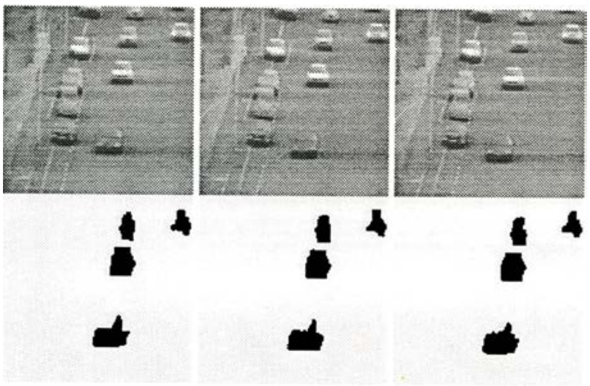
Part 1

IMAGE SEQUENCE ANALYSIS

2D Motion Detection



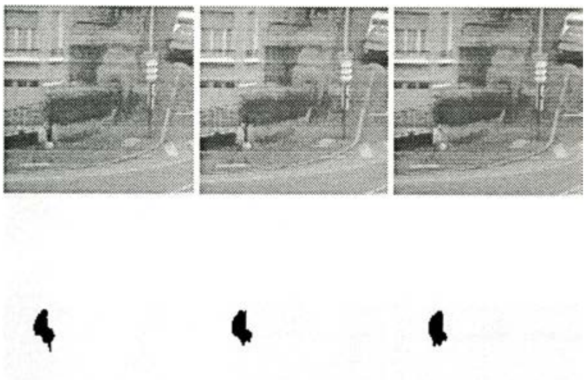
The 2D motion detection problem



P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- Situation :
  ▷ static camera, moving objects
  ▷ moving camera, moving objects (need of camera motion compensation)
- Issue : detection of moving objects in a scene
- Input : temporal changes in the image sequence
- Output : mobile object masks (binary information)
- Applications : road trafic control, electronic surveillance, trajectography.

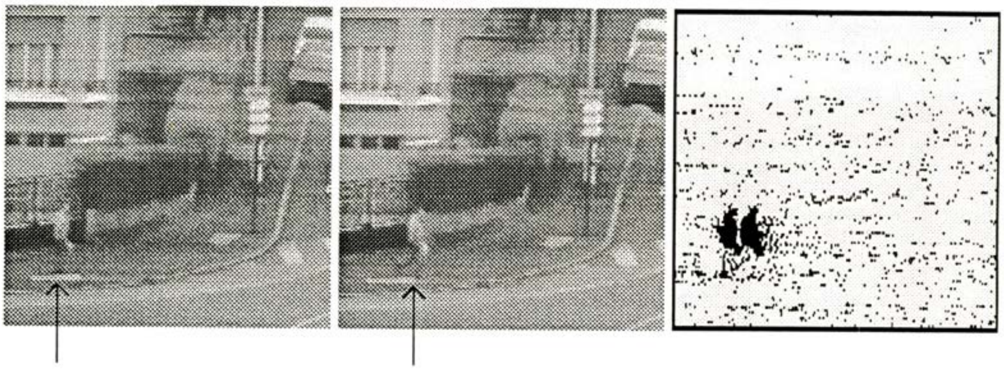
Hypotheses



P. Lalande, PHD Thesis, Univ. Rennes I, 1990

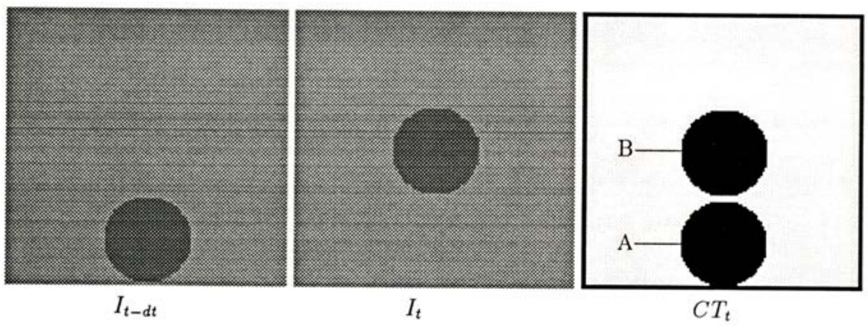
- Illumination of the scene is invariant
- Static camera
- No a priori knowledge on the motion of objects
- No a priori knowledge on the shape or luminosity of objects

Two problems to be solved

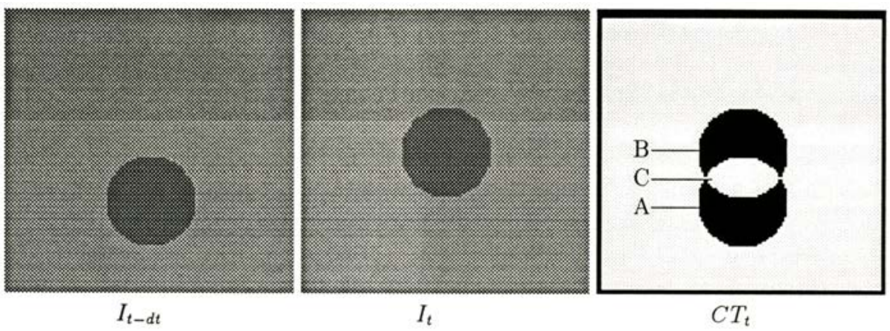


P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- Detection of temporal changes
- Reconstruction of the binary masks of moving objects



Temporal changes generated by a moving object (without overlap)

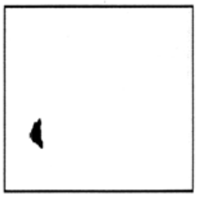


Temporal changes generated by a moving object (with overlap)

P. Lalande, PHD Thesis, Univ. Rennes I, 1990

Ecart entre les deux images: 1 seconde

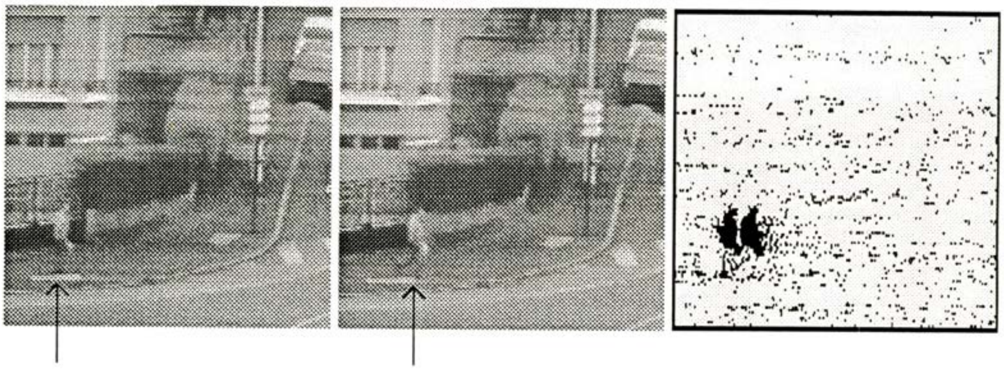
Ecart entre les deux images: 1/25^{è} de seconde



Temporal changes generated by a moving object (without and with overlap)

P. Lalande, PHD Thesis, Univ. Rennes I, 1990

Detection by frame difference (Jain78)



P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- Frame difference :

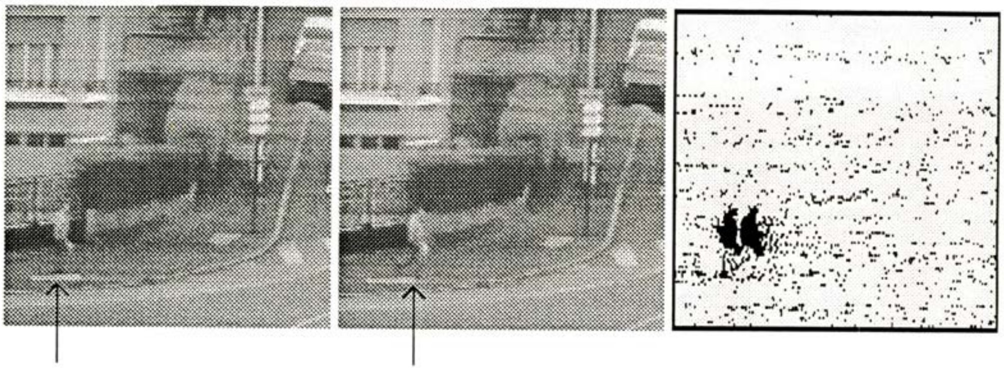
$$FD(x, y) = I_{t+dt}(x, y) - I_t(x, y)$$

- Detection of temporal changes :

$$CT(x, y) = \begin{cases} 255 & \text{if } |FD(x, y)| > \lambda \\ 0 & \text{else} \end{cases}$$

This method is very sensitive to noise.

Detection by averaging the frame difference (Jain78)



P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- Averaged frame difference :

$$NFD(x,y) = \frac{1}{N} \sum_{(u,v) \in V(x,y)} |I_{t+dt}(u,v) - I_t(u,v)|$$

  where $V(x,y)$ is a neighborhood of pixel (x,y) with N points.

- Detection of temporal changes :

$$CT(x,y) = \begin{cases} 255 & \text{if } NFD(x,y) > \lambda \\ 0 & \text{sinon} \end{cases}$$

This method remains sensitive to noise.

Detection by hypothesis test (Nagel84)

P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- **Model of image intensity :**

$$I(u,v) = \mu + n(u,v) \quad \forall (u,v) \in V(x,y)$$

  where : $V(x,y)$ is a square neighborhood of pixel (x,y) with N points,
  μ is a constant, $n(u,v)$ is a white gaussian noise $\sim \mathcal{N}(0,\sigma^2)$

- **Test of 2 hypotheses :**
  ▷ H_0 **(no temporal change) :**
  $\forall (u,v) \in V(x,y),\ I_t(x,y) \sim \mathcal{N}(\mu_0,\sigma^2)$ and $I_{t+dt}(x,y) \sim \mathcal{N}(\mu_0,\sigma^2)$

  ▷ H_1 **(temporal change) :**
  $\forall (u,v) \in V(x,y),\ I_t(x,y) \sim \mathcal{N}(\mu_1,\sigma^2)$ and $I_{t+dt}(x,y) \sim \mathcal{N}(\mu_2,\sigma^2)$ with $\mu_1 \neq \mu_2$

Detection by hypothesis test (Nagel84)

P. Lalande, PHD Thesis, Univ. Rennes I, 1990

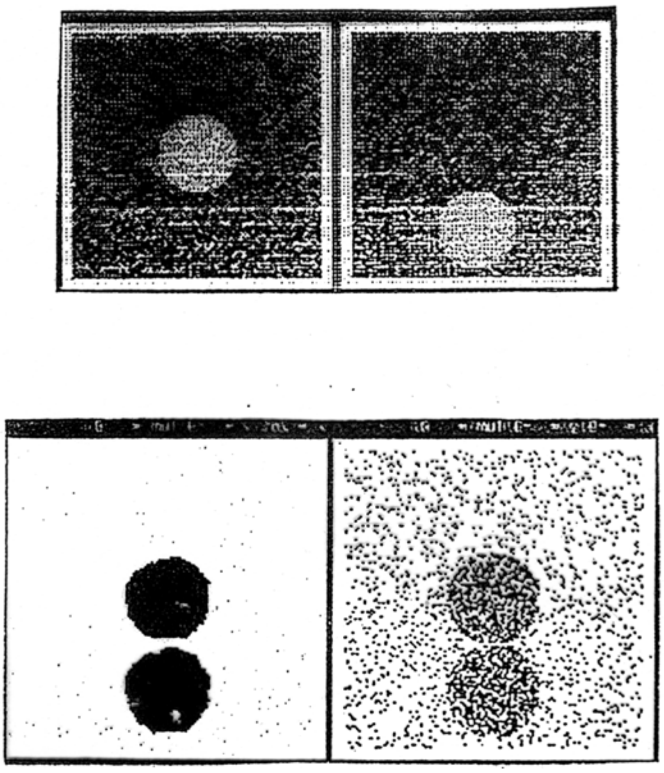
- **Generalized likelihood ratio test (GLRT) :**

$$\max_{\mu_1,\mu_2} \min_{\mu_0} \frac{p\left(\left\{I_t(u,v), I_{t+dt}(u,v)\ ,\ (u,v)\in V(x,y)\right\}\ \mid\ H_1\right)}{p\left(\left\{I_t(u,v), \{I_{t+dt}(u,v)\ ,\ (u,v)\in V(x,y)\right\}\ \mid\ H_0\right)} \begin{array}{c}(H_1)\\ >\\ <\\ (H_0)\end{array} \lambda$$

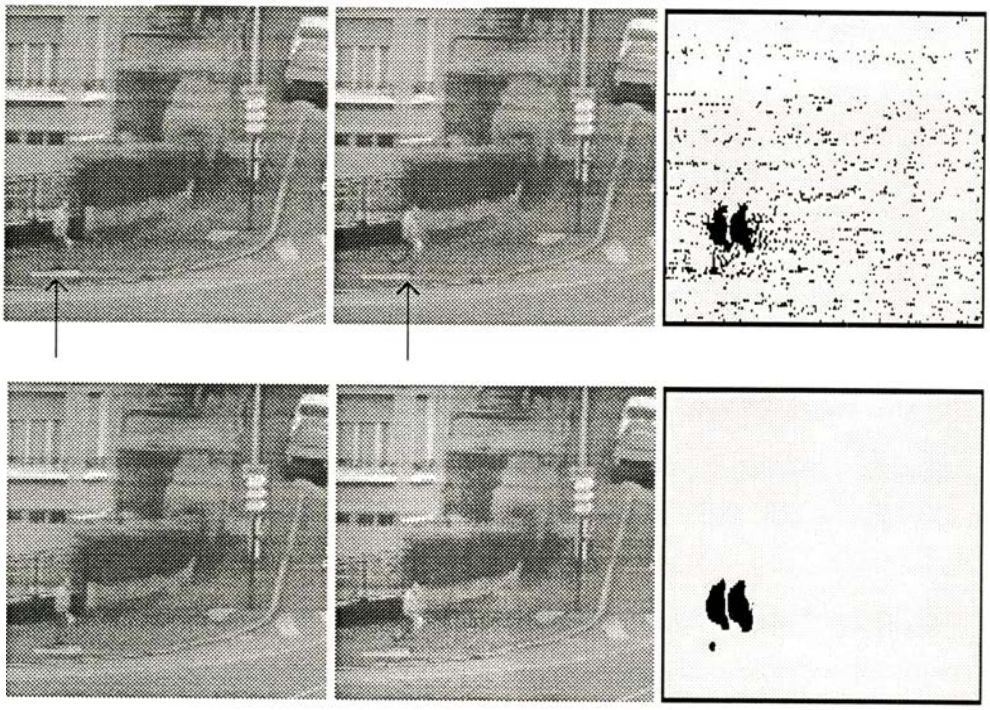
$$\Leftrightarrow \quad R(x,y) = \frac{1}{2\sigma\sqrt{N}} \left|\ \sum_{(u,v)\in V(x,y)} I_{t+dt}(u,v) - I_t(u,v)\ \right| \begin{array}{c}(H_1)\\ >\\ <\\ (H_0)\end{array} \lambda'$$

- **Change detection :**

$$CT(x,y) = \begin{cases} 255 & \text{if } R(x,y) \geq \lambda' \\ 0 & \text{else} \end{cases}$$



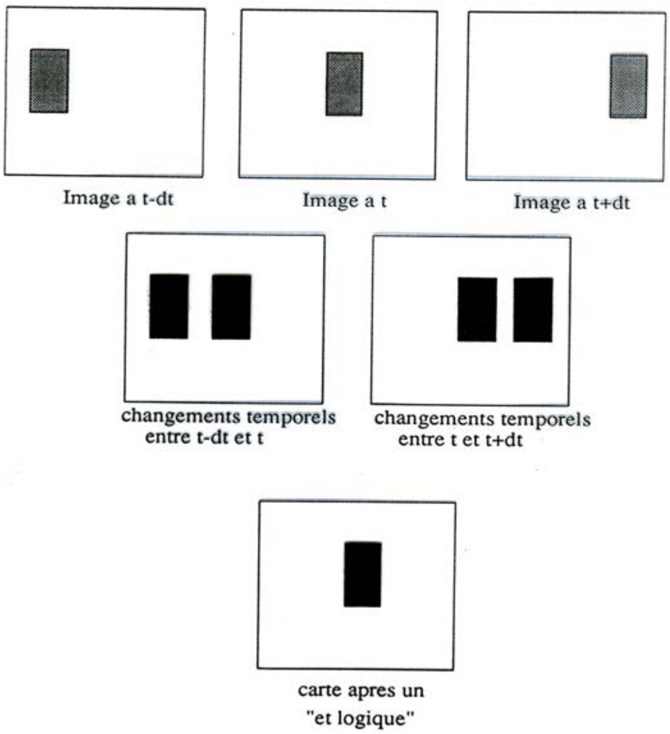
P. Lalande, PHD Thesis, Univ. Rennes I, 1990

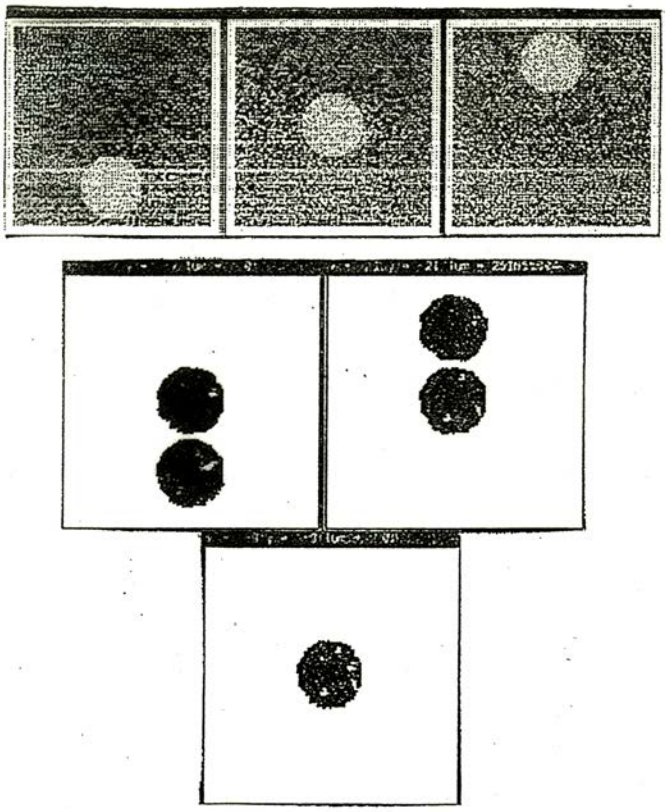


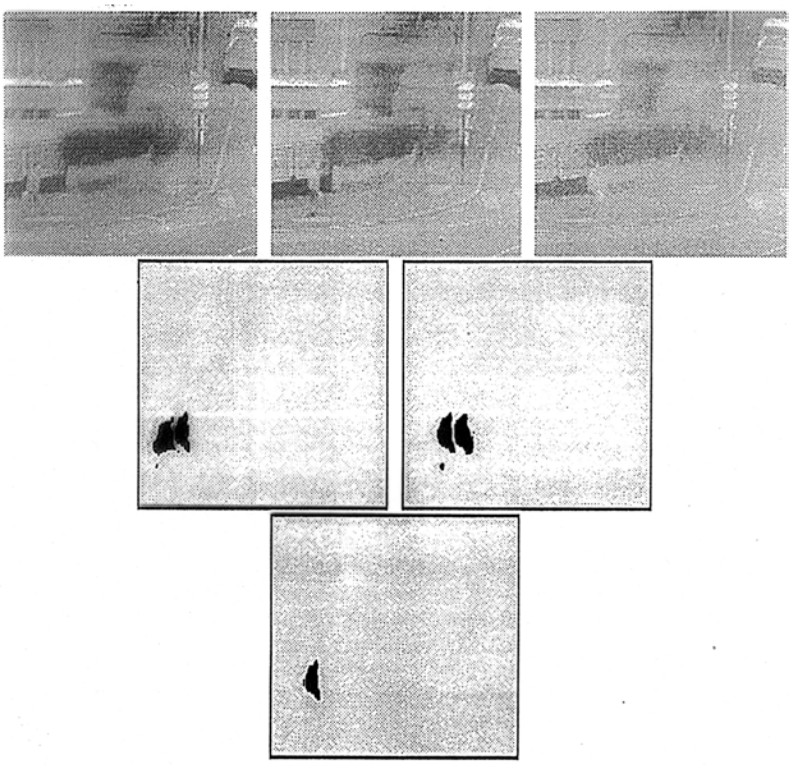
P. Lalande, PHD Thesis, Univ. Rennes I, 1990

Reconstruction of the mask

Using a binary “AND" operator (Jain, Martin, Aggarwal79)

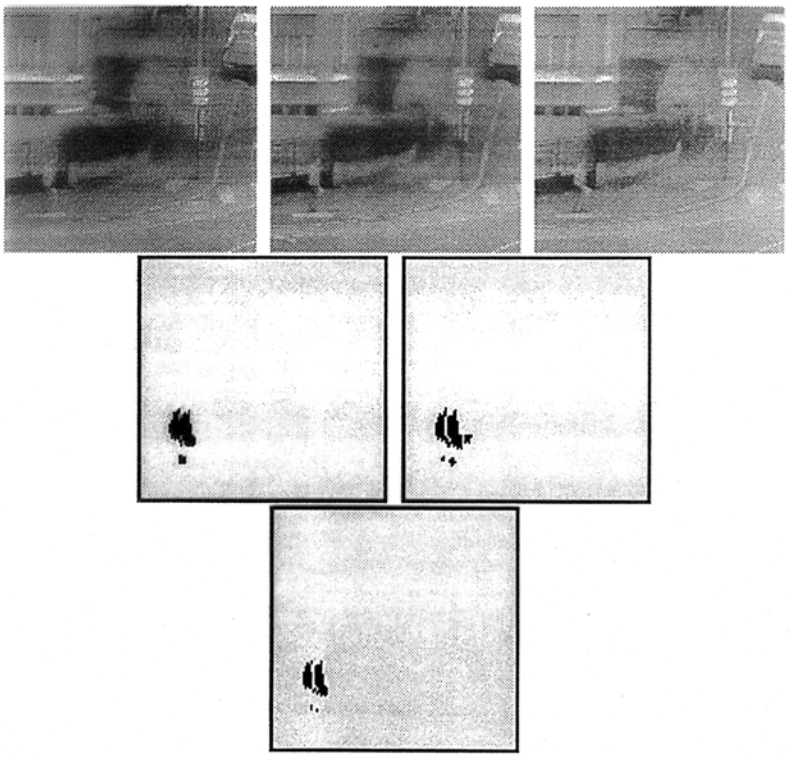






P. Lalande, PHD Thesis, Univ. Rennes I, 1990

Time difference between the 2 images : 1 s



P. Lalande, PHD Thesis, Univ. Rennes I, 1990

Time difference between the 2 images : 1/25 s

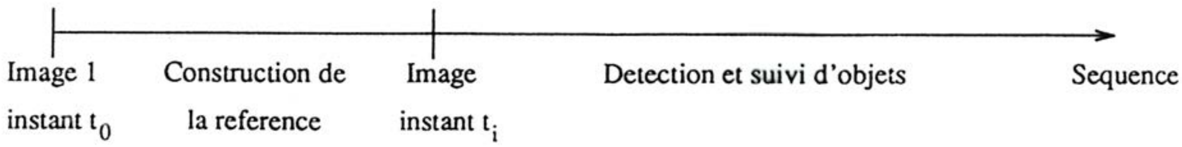
Background subtraction techniques

reference image current image change detection

P. Lalande, PHD Thesis, Univ. Rennes I, 1990

- **Principle :**
  - ▷ Construction of a reference image (background)
  - ▷ Detection of changes between current image and reference image
  - ▷ Updating of reference image

Reference image construction and updating



Simple algorithm example

$p = (x, y)$, $I(p, t_n)$ = image at time t_n, $I_{ref}(p, t_n)$ = reference (background) image at time t_n

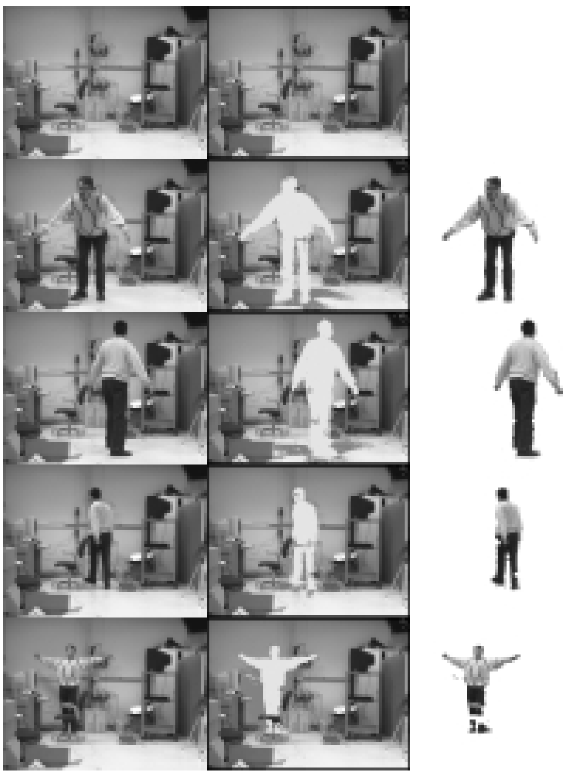
- **Reference image construction :** $I_{ref}(p, t_n) \ = \ I(p, t_n) \ + \alpha\,(I_{ref}(p, t_{n-1}) \ - \ I(p, t_n))$

with :

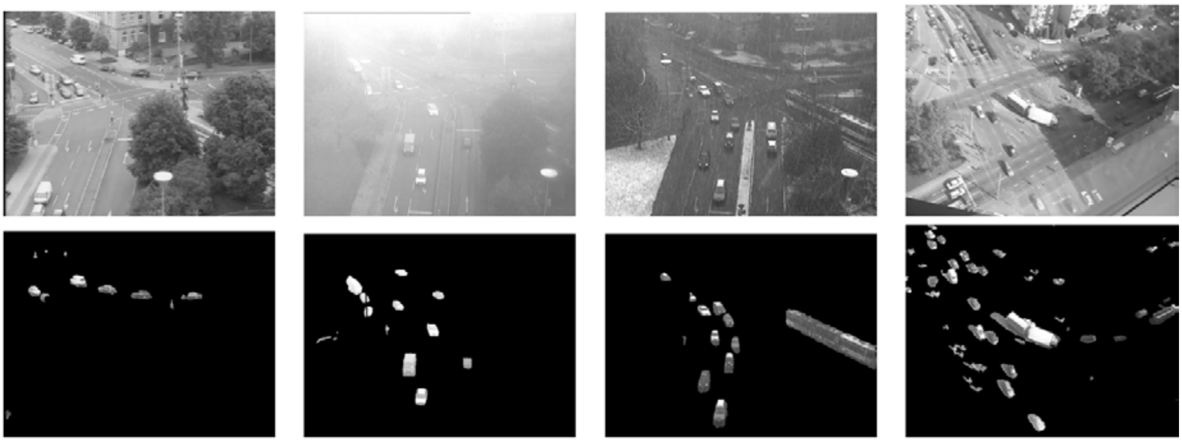
▷ $I_{ref}(p, t_0) = NIL \quad \forall p \in I_{ref}$

▷ $\alpha = \begin{cases} 0 & \text{if } I_{ref}(p, t_{n-1}) = NIL \text{ and no temporal change detected at p} \\ 0 \leq \alpha \leq 1 & \text{if } I_{ref}(p, t_{n-1}) \neq NIL \text{ and no temporal change detected at p} \\ 1 & \text{if temporal change detected at p} \end{cases}$

- **Reference image updating :** $I_{ref}(p, t_n) \ = \ I(p, t_n) \ + \alpha\,(I_{ref}(p, t_{n-1}) \ - \ I(p, t_n))$



T. Horprasert, D. Harwood, L.S.Davis, A Robust Background Subtraction and Shadow Detection, ACCV, 2000



S. Cheung, C. Kamath, Robust techniques for background subtraction in urban traffic video, Proc. SPIE 5308, 881, 2004

Télécom Physique Strasbourg

Module 36032

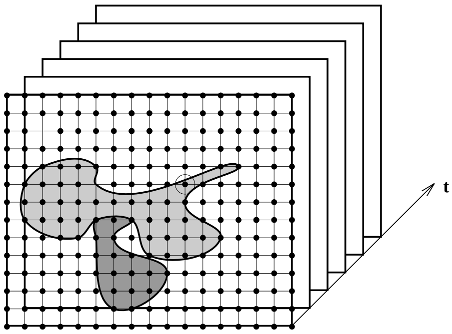
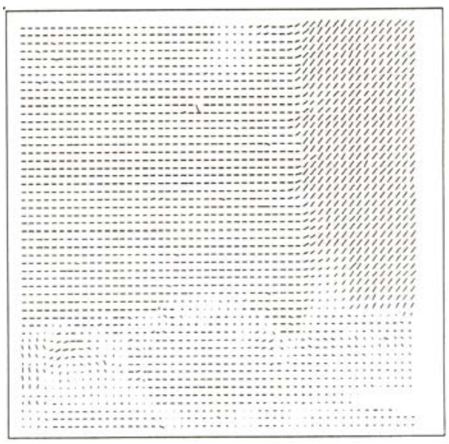


IMAGE SEQUENCE ANALYSIS

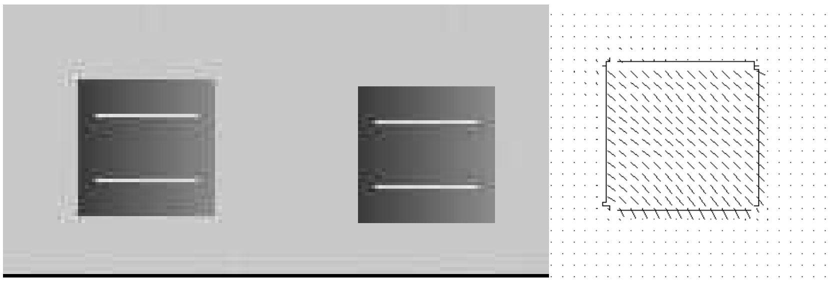
Part 2

IMAGE SEQUENCE ANALYSIS

2D Motion Estimation

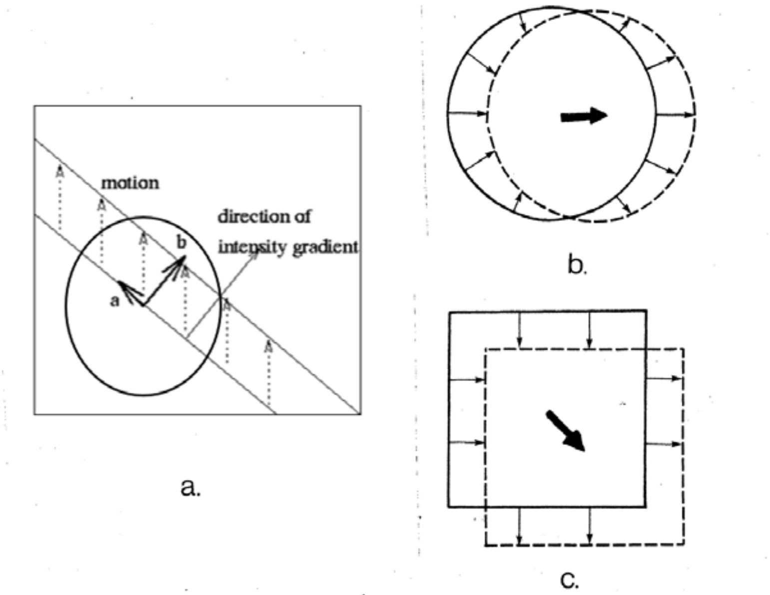


The 2D motion estimation problem



- Estimation (measurement) of motion in the 2D image plane
- Ambiguities in motion perception and estimation
- Input : spatiotemporal changes in the image sequence
- Output : displacement or velocity vectors
- Applications : scene analysis, interframe coding and compression (MPEG)

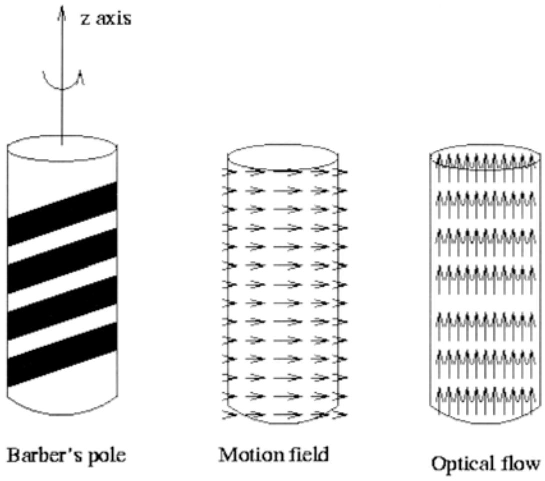
Motion ambiguities : the aperture problem



E.C. Hildreth, MIT AI Memo NO 734, Sept. 1983, E. Bichot, 2002

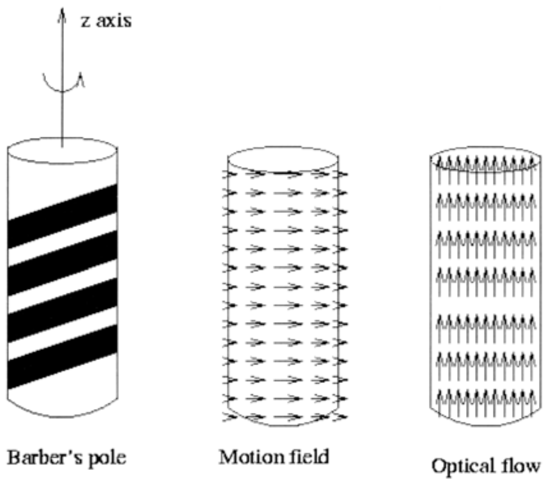
a) An operation that views a moving edge E through a local aperture. One can perceive (and compute) only the component of motion perpendicular to the edge. b) and c) The perpendicular components of velocity for a translating circle and square.

Motion ambiguities : the aperture problem



The barberpole example

Optical flow and motion field



- Motion field : projection of the 3D velocities of points on the 2D image plane.
  ▷ induced by the relative motion between camera and scene.
  ▷ not completely observable or measurable
- Optical flow : apparent motion of the image brightness pattern
  ▷ this is what we measure

Optical flow and motion field

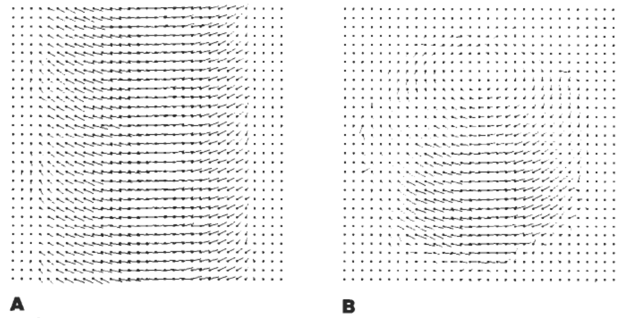


FIG. 10. Flow patterns computed for a cylinder rotating about its axis and for a rotating sphere. The axis of the cylinder is inclined 30 degrees towards the viewer and that of the sphere 45 degrees. Both are rotating at about 5 degrees per time step. The estimates shown are obtained after 32 time steps.

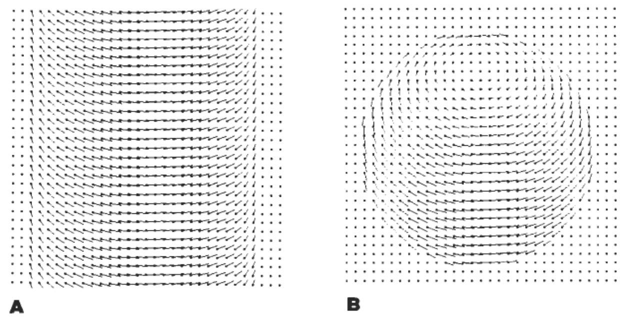
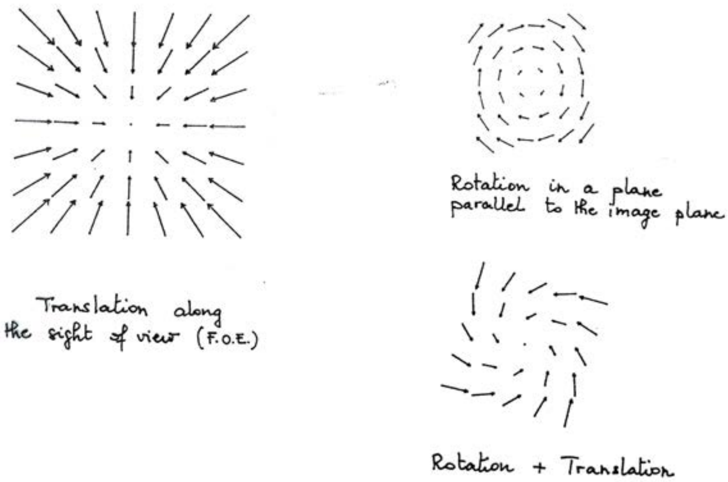


FIG. 11. Exact flow patterns for the cylinder and the sphere.

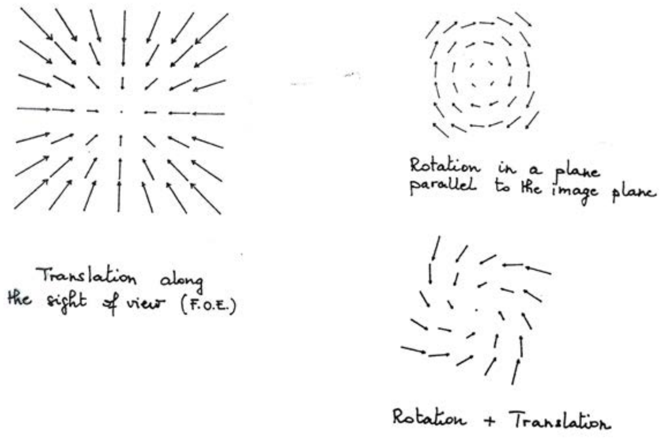
B.K.P. Horn, B.G. Schunck, Determining Optical Flow, Artificial Intelligence, Vol. 17, pp. 185-203, 1981

Optical flow estimation



- Not directly measurable
- Representation :
  ▷ velocity vector : $\omega \ = \ (\frac{dx}{dt}, \frac{dy}{dt})$
  ▷ displacement vector : $d \ = \ \omega.\Delta t$

Optical flow estimation



- Approaches :
  - ▷ Image features matching
  - ▷ Differential methods
  - ▷ Spatio-temporal filtering
- Different kind of measurements :
  - ▷ dense / sparse optical flow fields
  - ▷ local / global measurements
  - ▷ restrictions on the type of motion

2D Motion estimation

Image features matching

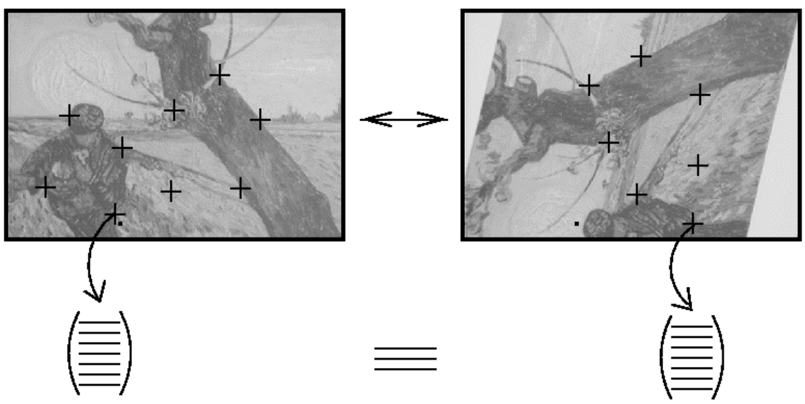
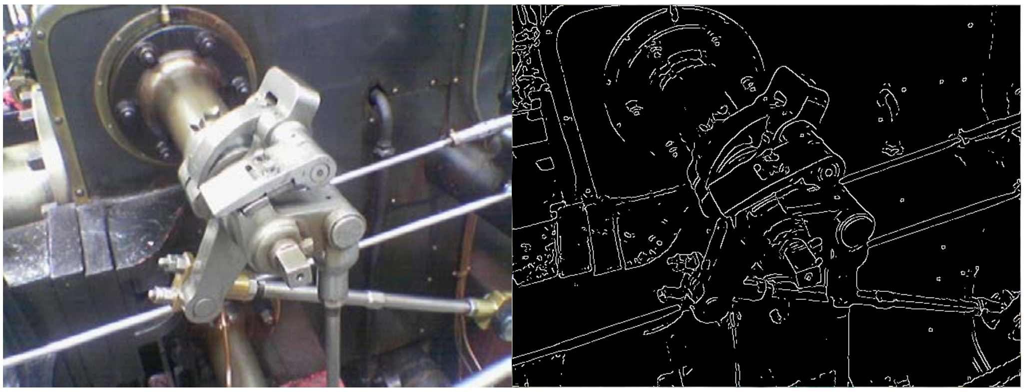
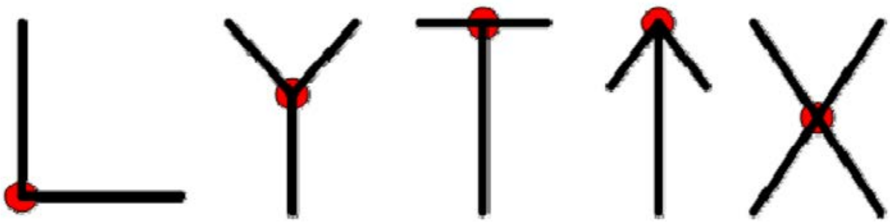


Image features



- Raw pixels
- Blocks of pixels
- Points of interest (corners, Harris, Moravec points, junctions, SIFT)
- Edge points, contour segments
- Segmented regions
- Graphs

Points of interest : corners, junctions

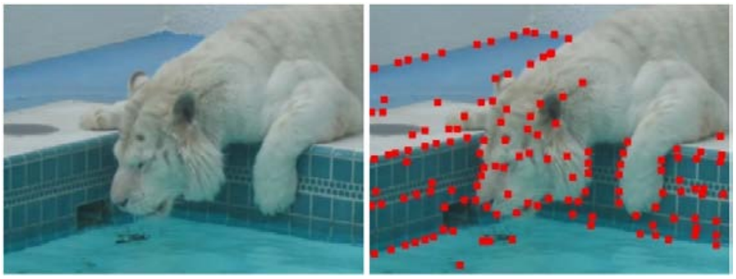


Corner=L-junction, Y-junction, T-junction, Arrow-junction, and X-junction (from : D.Parks, J.P. Gravel, Corner Detection)

Real corner (left) and Y-junction (right)

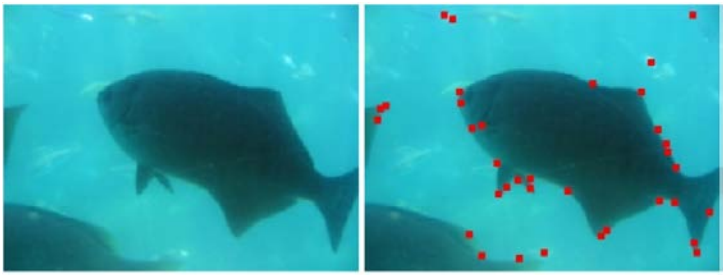
(M.A. Ruzon, C. Tomasi, Edge, junction and corner detection using color distributions, IEEE Trans. PAMI, Vol. 23, Nov. 2001).

Corner detection



©Robot Realm http ://www.roborealm.com/help/Moravec.php

Moravec corners (1980)



©Robot Realm http ://www.roborealm.com/help/Harris.php

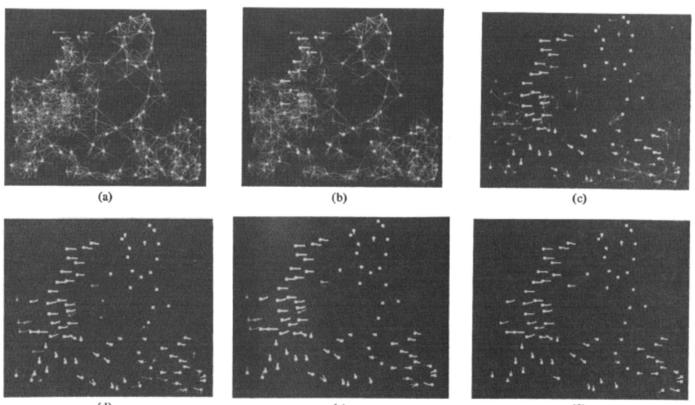
Harris corners (1988)

Moravec corners matching

BARNARD AND THOMPSON: DISPARITY ANALYSIS OF IMAGES

(a) (b)

Fig. 1. Stereogram.

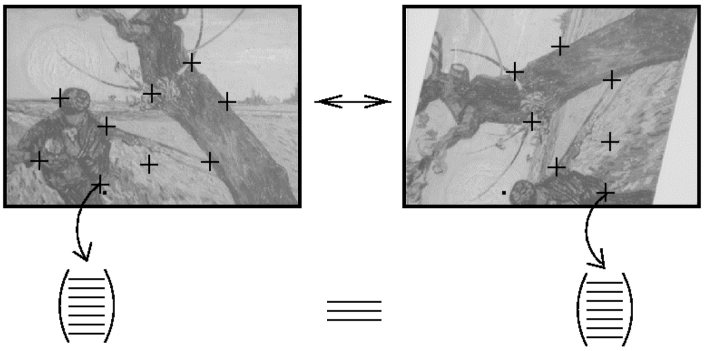


(a) (b) (c) (d) (e) (f)

Fig. 2. (a) Initial probability assignments for Fig. 1. (b)–(f) Iterations 2, 4, 6, 8, and 10.

S. Barnard, W. Thompson, Disparity Analysis of Images, IEEE Trans. PAMI, July 1980

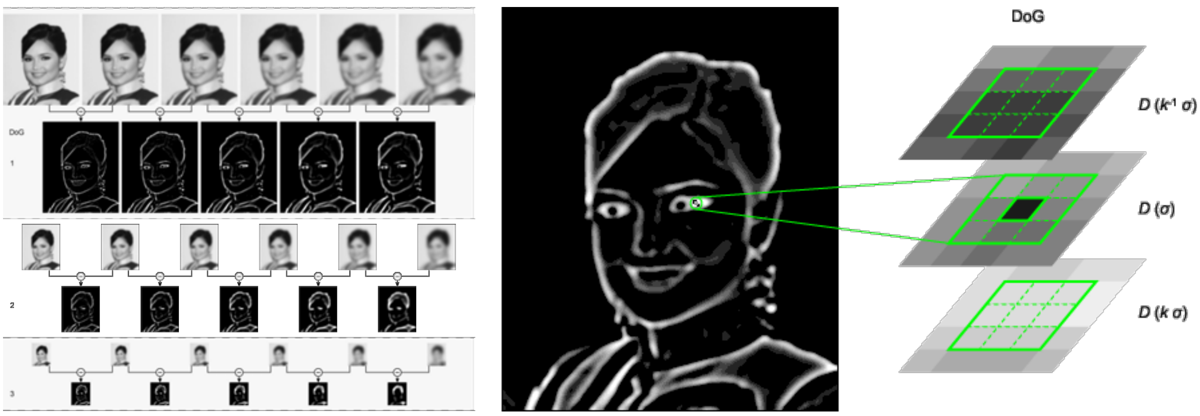
SIFT : Scale Invariant Feature Transform (D. Lowe 1999)



C. Schmid, Movi, INRIA Rhône-Alpes

- Gaussian pyramids, Difference of gaussians
- Detection of extrema in scale space
- Local histograms of orientations = SIFT descriptors (vectors size : 128)
- SIFT points are invariant to translation, rotation and scale !

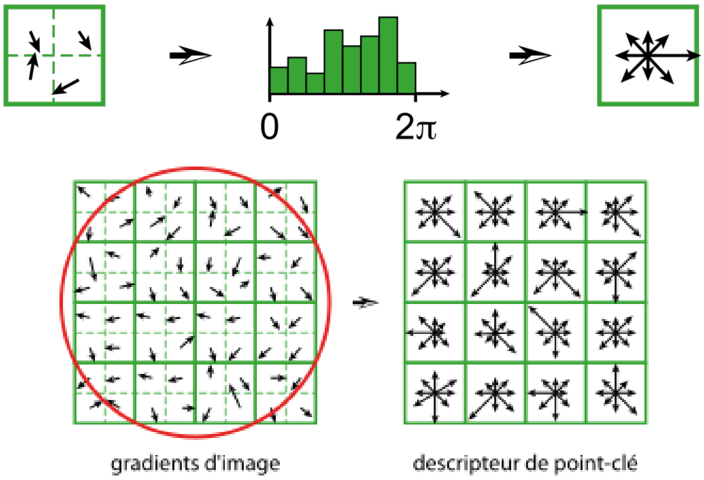
SIFT : Scale Invariant Feature Transform (D. Lowe 1999)



Wikipedia © amrfum

- Gaussian pyramids, Difference of gaussians
- Detection of extrema points on DoG images (26-neighborhood)
- $\Rightarrow$ SIFT point = (x, y, σ)

SIFT : Scale Invariant Feature Transform (D. Lowe 1999)



Wikipedia © amrfum

- Local histograms of gradient orientations at extrema points (in a neighborhood)
- Main orientation (histogram mode) = θ. Definition of a local frame (rotation of angle θ)
- SIFT point = (x, y, σ, Θ) + Descriptor = orientations histogram (16x16 neighborhood, 16 histograms, 8 directions = vector size 128)

Points of interest matching : example

(a) Scale change of 3.9 and rotation of 17°.

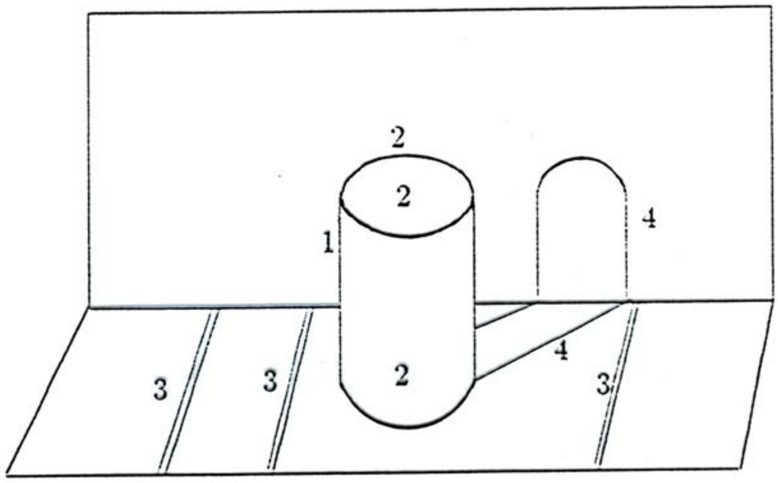
(b) Scale change of 1.8 and viewpoint change of 30°

(c) Scale change of 1.7 and viewpoint change of 50°

K. Mikolajczyk, C. Schmid, Scale and Affine Invariant Interest Point Detectors,

Int. Jl. Computer Vision, Vol. 60, No 1, 2004

Edge features



N. Ayache, Vision stéréoscopique et perception multisensorielle : application à la robotique, Masson, 1989

Different kind of edges :

1. Occlusions
2. Ridges
3. Textures
4. Shadows

Edge matching

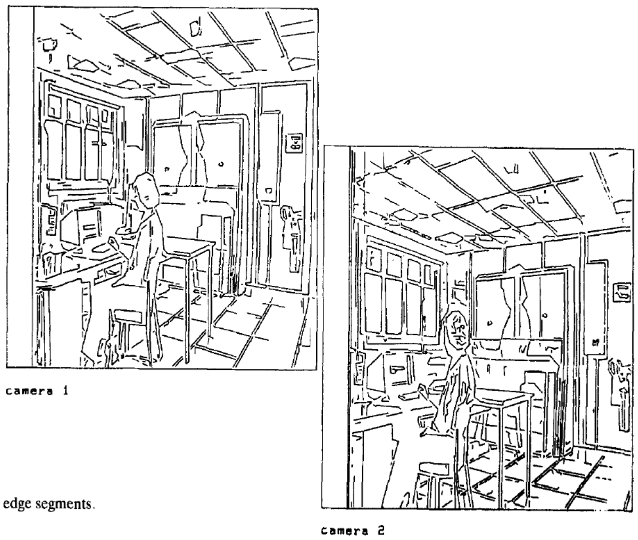


Fig. 9. Initial edge segments.

N. Ayache, B. Faverjon, Efficient Registration of Stereo Images by Matching Graph Descriptions of Edge Segments, Int. Jl. Computer Vision, Vol. 1, No 2, 1987

Initial edge segments

Edge matching

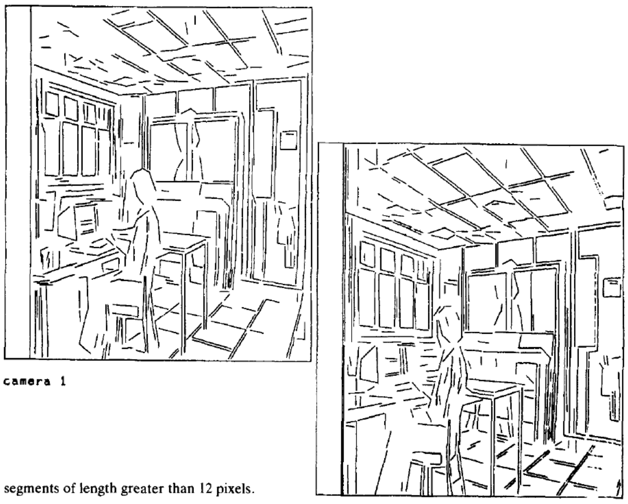


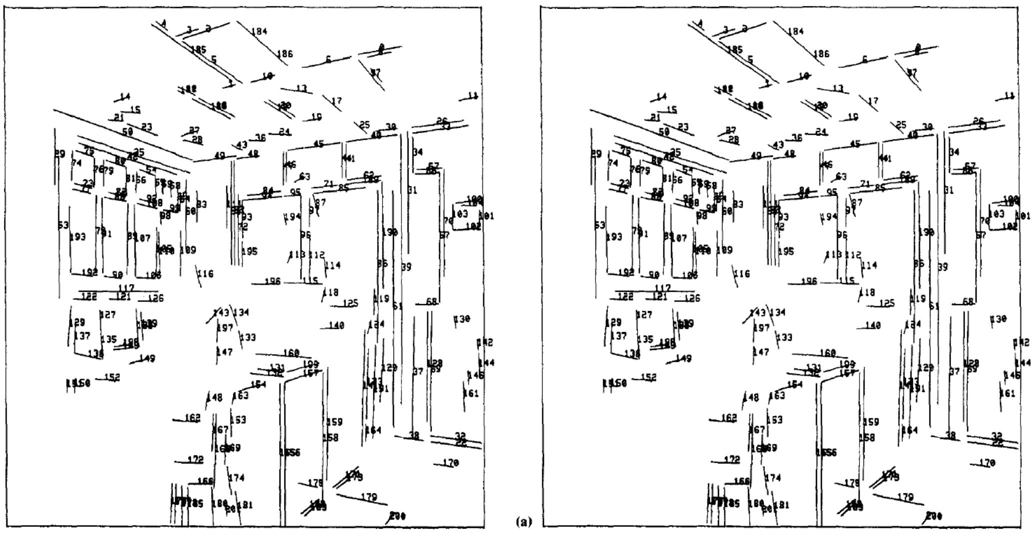
Fig. 10. Edge segments of length greater than 12 pixels.

N. Ayache, B. Faverjon, Efficient Registration of Stereo Images by Matching Graph Descriptions of Edge Segments, Int. Jl. Computer Vision, Vol. 1, No 2, 1987

Edge segments of length greater than 12 pixels

Edge matching

(a) (a)

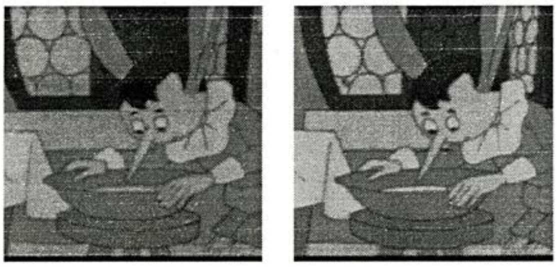


N. Ayache, B. Faverjon, Efficient Registration of Stereo Images by Matching Graph Descriptions of Edge Segments, Int. Jl. Computer Vision, Vol. 1, No 2, 1987

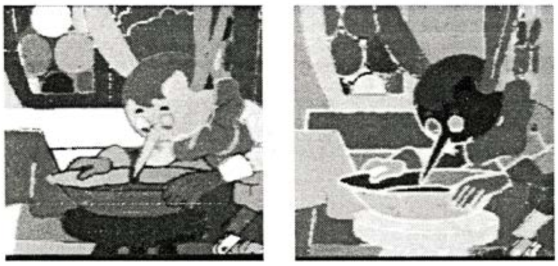
Matched segments between camera 1 and camera 2

Regions

images originales



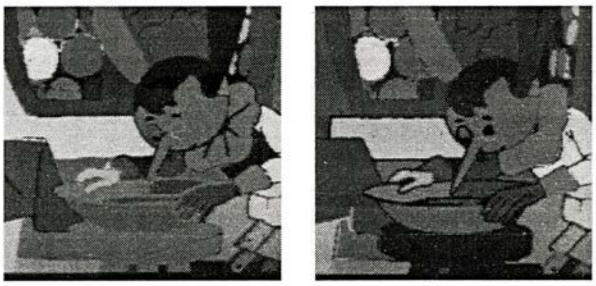
images segmentées



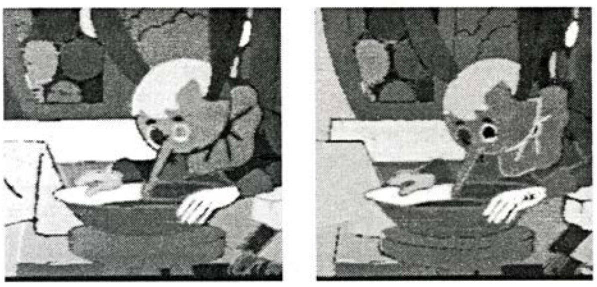
Y. Remion, Stéréovision par zones. Outils et structure d'un système expert, Thèse, Télécom Paris, 1988

Region matching

recalage : critère n° 1



recalage : critère n° 2



Y. Remion, Stéréovision par zones. Outils et structure d'un système expert, Thèse, Télécom Paris, 1988

Region matching

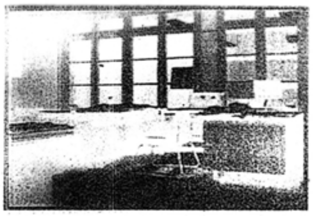


Fig. 11. – Image originale gauche.

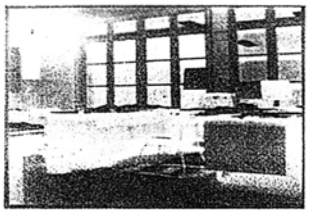


Fig. 12. – Image originale droite.

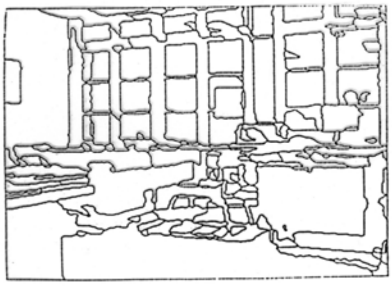


Fig. 13. – Segmentation finale (algorithme n° 2) (image gauche).

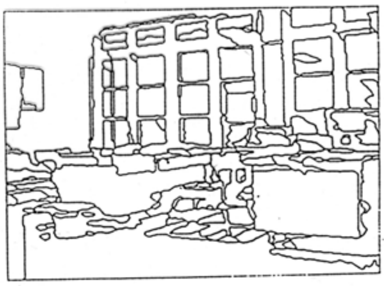


Fig. 14. – Segmentation finale (algorithme n° 2). (image droite).

O. Monga, B. Wrobel, Segmentation d'images : vers une méthodologie, Revue Traitement du Signal, Vol. 4, No 3, 1987

Image representation by graphs

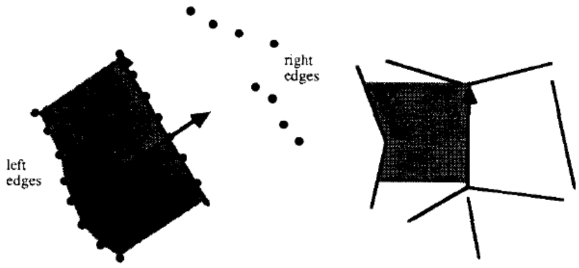


Fig. 1. A line and the detection of its adjoining regions (left), and the final local configuration associated with a line (right).

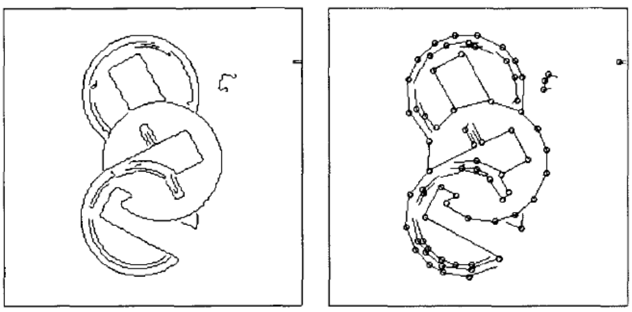


Fig. 2. Edges extracted from curved objects (left) and the associated structural description (right). The *left_of* and *right_of* relations are not shown.

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques,

IEEE Trans. PAMI, Vol. 11, No 11, 1989

Image representation by graphs

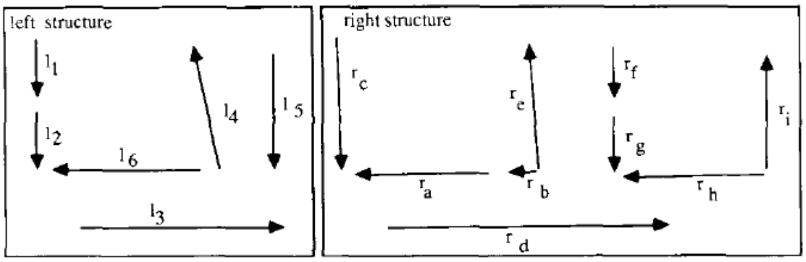


Fig. 5. Two images to be matched.

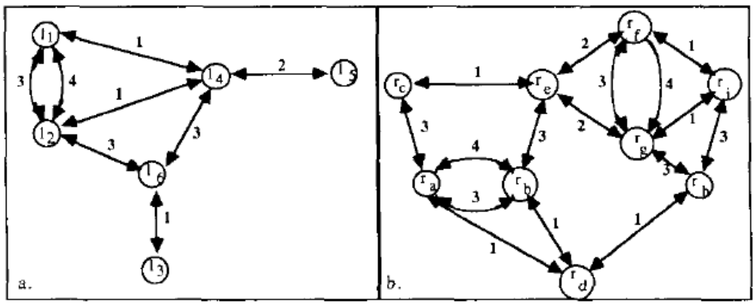


Fig. 6. Two structural descriptions to be matched. The interline relations are: left of (1), right of (2), same junction (3), and collinear (4).

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques,

IEEE Trans. PAMI, Vol. 11, No 11, 1989

Graph matching

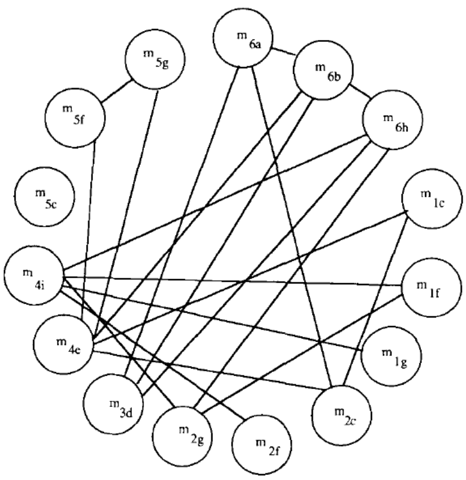


Fig. 7. The graph representation after node building and after applying the first three rules.

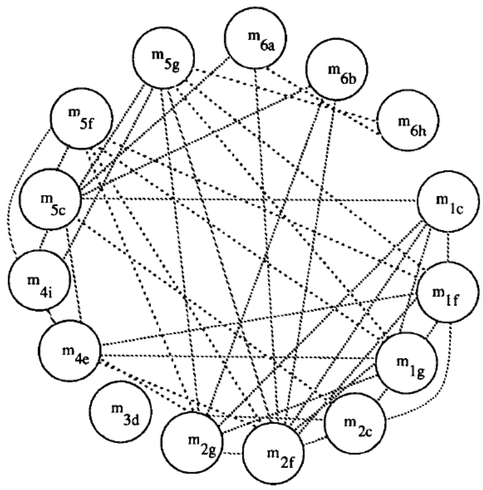


Fig. 8. The incompatible nodes found after applying the fourth rule.

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques,

IEEE Trans. PAMI, Vol. 11, No 11, 1989

Graph matching

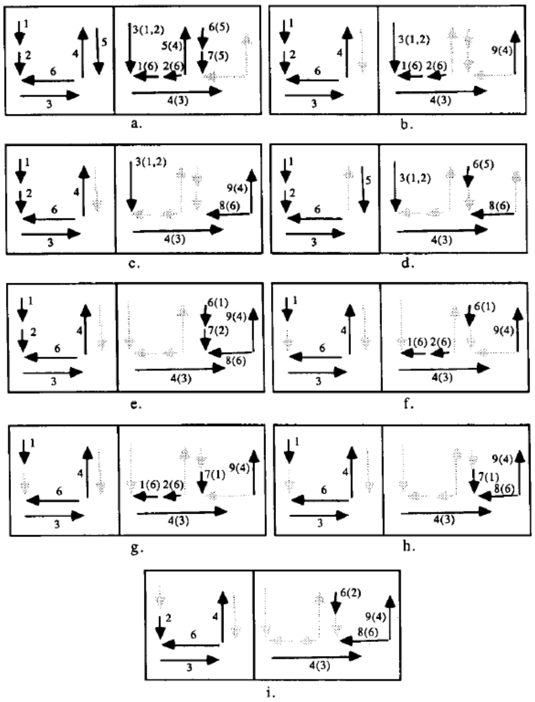


Fig. 9. All the matches corresponding to nine maximal cliques. The best solution (a) is the unique largest maximal clique found in the graph.

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques,

IEEE Trans. PAMI, Vol. 11, No 11, 1989

Graph matching

Fig. 10. The first image pair. The three windows on which the local matching is demonstrated are shown on the left image.

Fig. 11. Lines extracted from the above images.

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques,

IEEE Trans. PAMI, Vol. 11, No 11, 1989

Graph matching

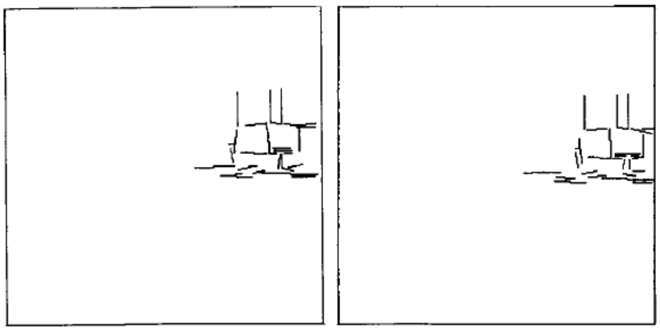


Fig. 13. First example: the result of matching (35 lines).

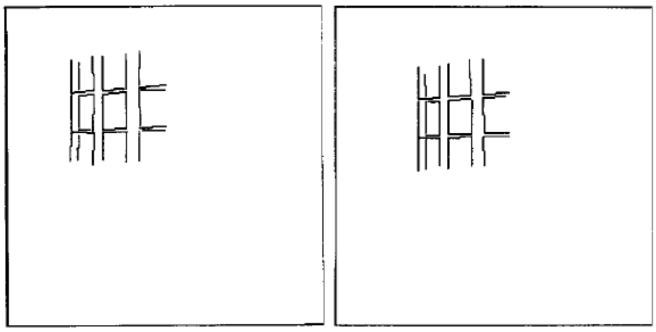


Fig. 14. Second example: the result of matching (33 lines).

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques, IEEE Trans. PAMI, Vol. 11, No 11, 1989

Graph matching

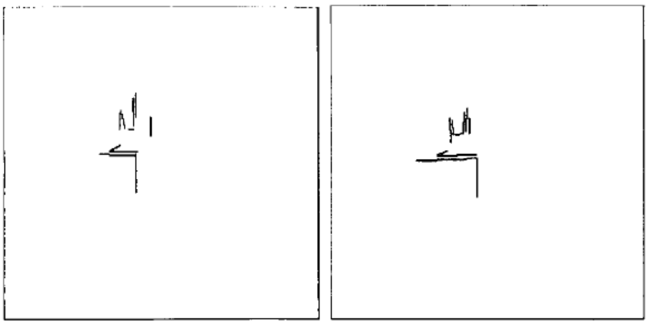


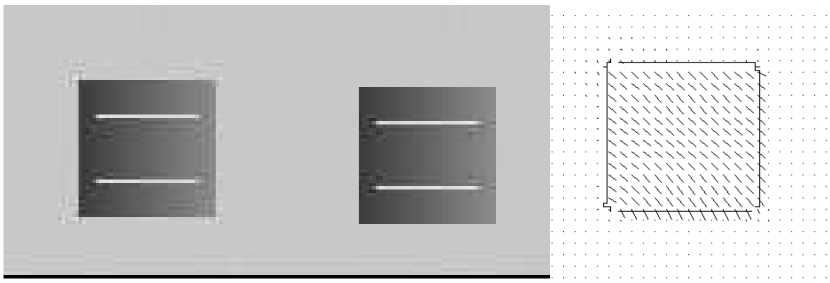
Fig. 15. Third example: the result of matching (10 lines).

Fig. 16. The matched lines for the first image pair.

R. Horaud, T. Skordas, Stereo Correspondence Through Feature Grouping and Maximal Cliques, IEEE Trans. PAMI, Vol. 11, No 11, 1989

2D Motion estimation

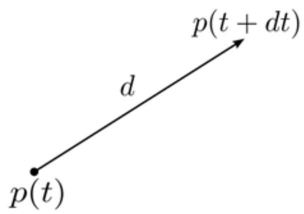
Differential approaches



Displaced Frame Difference (DFD)

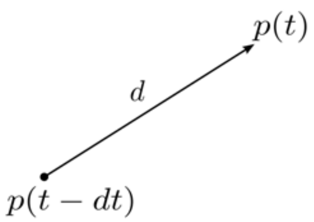
- f : intensity function
- moving point : P. At time t, P is at position $p = (x, y)$
- displacement : $d = (dx, dy)$
- DFD = displaced frame difference (différence interimage déplacée)

- definition 1 : displacement from t to $t + dt$



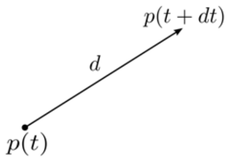
$$DFD(p, d, t) \;=\; f(p + d, t + dt) - f(p, t)$$

- definition 2 : displacement from $t - dt$ to t



$$DFD(p, d, t) \;=\; f(p, t) - f(p - d, t - dt)$$

Differential approach



Basic assumption : the brightness of a particular moving point is constant over time.

• If point $P(p = (x, y))$ moves according to $d = (dx, dy)$ from time t to time $t + dt$:
$f(p + d, t + dt) \;=\; f(p, t)$

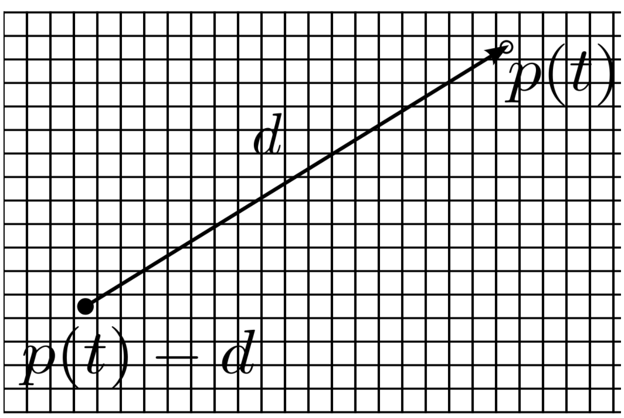
• The estimation of the displacement d amounts to solve the following equation :
$DFD(p, d, t) \;=\; 0$
or (because of noise) :
$\hat{d} = \arg\min_{\boldsymbol{d}} \|DFD(p, d, t)\|$ where $\| \,.\, \|$ is some norm.

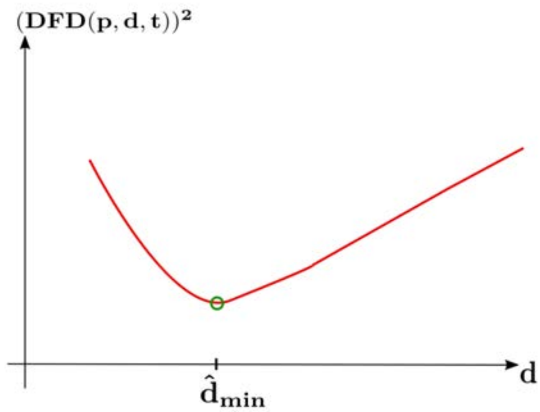
• Implications of the constant brightness assumption are :
▷ illumination must be constant over time
▷ shading effects are ignored
▷ no brightness variation due to variations in 3D surface orientations

• Differential techniques differ depending on :
▷ how the DFD is considered (definition, norm)

▷ which minimization procedure is chosen

Method 1 : pel-recursive method





$\mathbf{DFD(p, d, t) = f(p, t) - f(p - d, t - dt)}$

• Direct minimization of $(DFD(p, d, t))^{\mathbf{2}}$ (Netravali-Robbins 1979) :
$\hat{d} = \arg\min_{\boldsymbol{d}} \, (DFD(p, d, t))^{\mathbf{2}}$

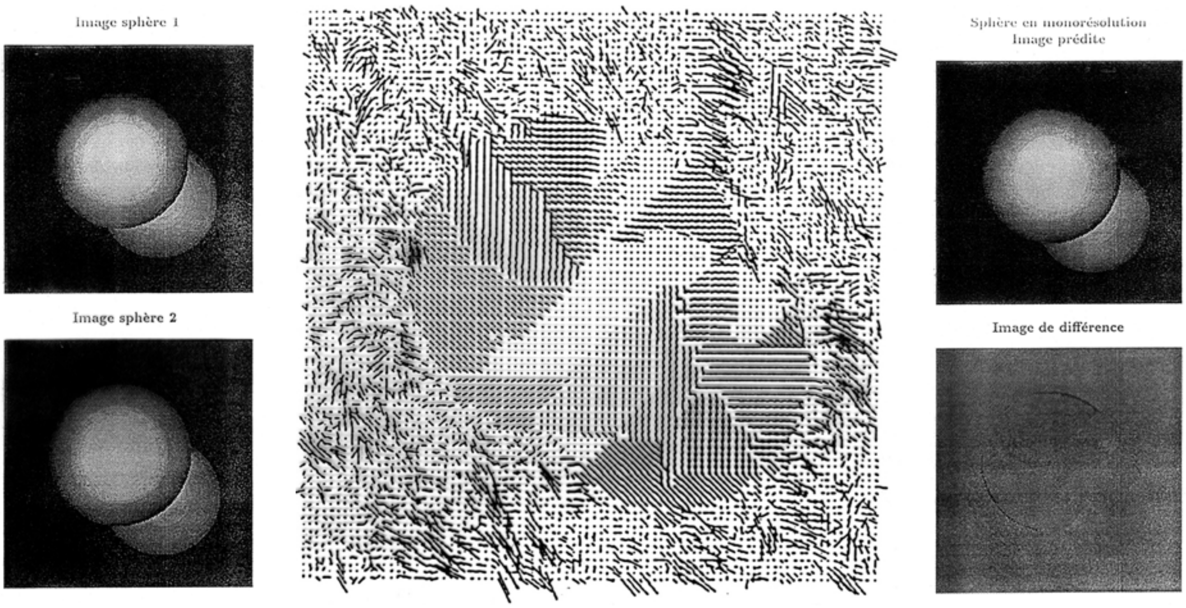
• A steepest descent algorithm is used (fixed step size) :

$$\hat{d}(i) = \hat{d}(i-1) \;-\; \tfrac{\epsilon}{2} \, \nabla_{\boldsymbol{d}} \, (DFD(p, d, t))^{\mathbf{2}} \Big|_{\boldsymbol{d} = \hat{\boldsymbol{d}}(\boldsymbol{i}-\mathbf{1})}$$

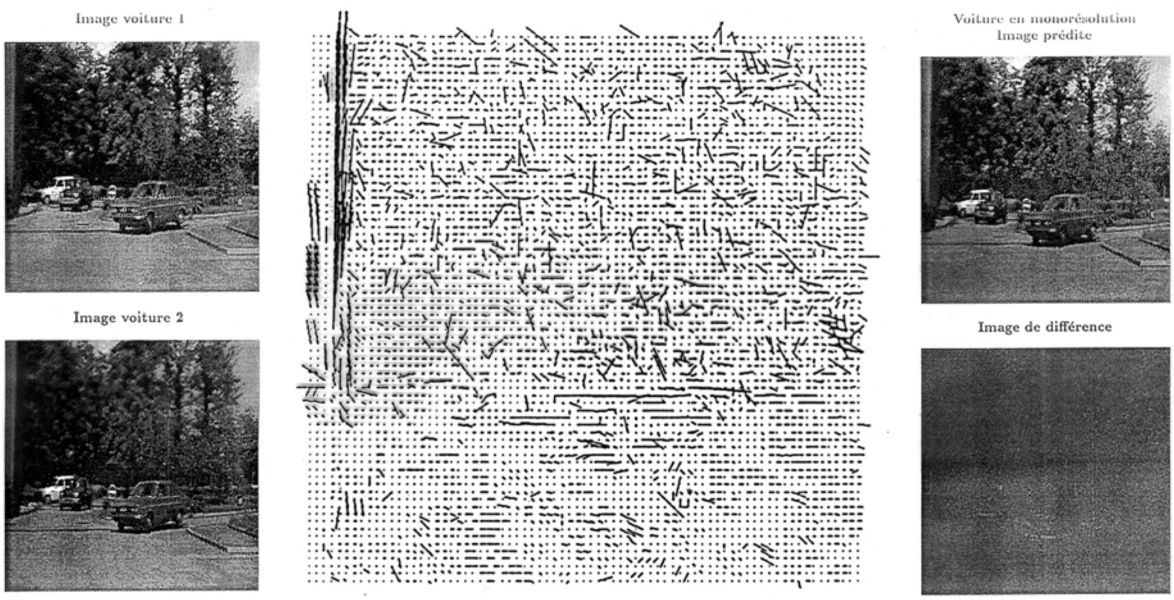
$$\Leftrightarrow \;\; \hat{d}(i) = \hat{d}(i-1) \;-\; \epsilon \, DFD(p, \hat{d}(i-1), t) \, \nabla f(p - \hat{d}(i-1), t - dt)$$

• In practice, recursion is carried out from one pixel $p_{\boldsymbol{k}}$ to its neighbor $p_{\boldsymbol{k}+\mathbf{1}}$ along a line of the image.

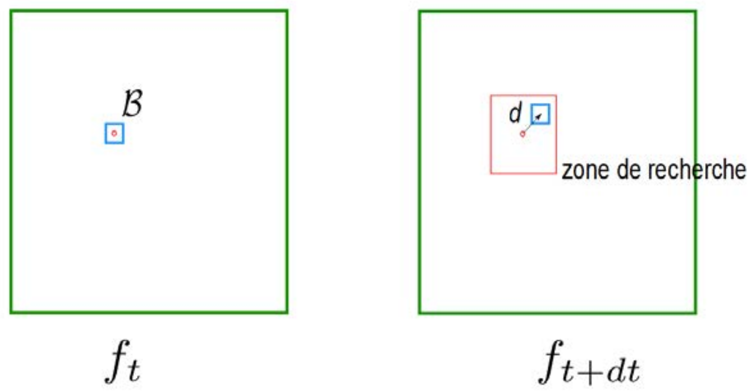
Method 1 : pel-recursive method



Method 1 : pel-recursive method



Method 2 : block matching method

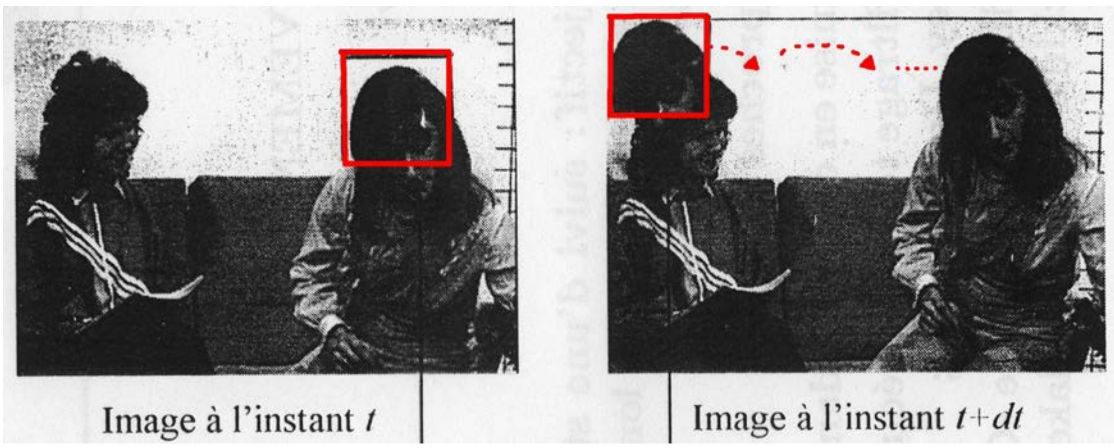


• Displacement d is assumed to be constant within a block $\mathcal{B}$, around the current pixel p
$DFD(p, d, t) \ = \ f(p + d, t + dt) - f(p, t)$

$$\hat{d} = \arg\min_{\boldsymbol{d}} \sum_{\boldsymbol{p_k} \in \boldsymbol{B}} \|DFD(p_k, d, t)\|$$

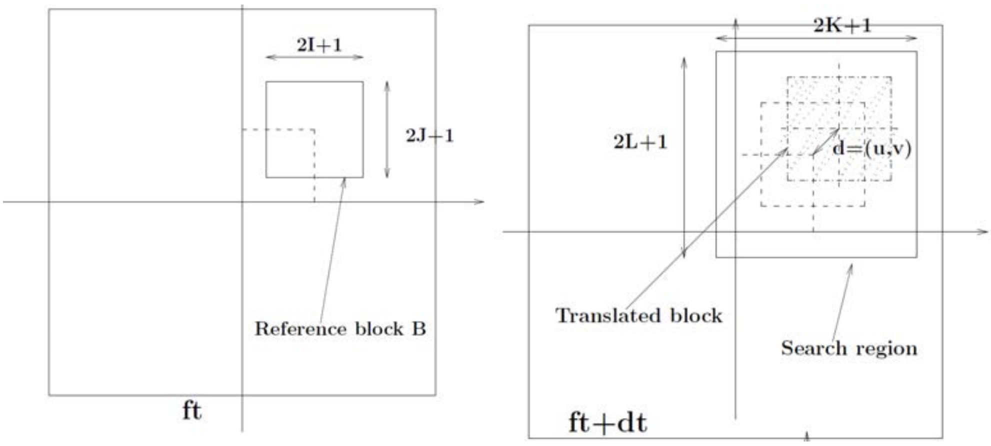
where $\| \, . \, \|$ is the $(.)^2$ or $| \, . \, |$ norm.

Method 2 : block matching method



IRISA/INRIA Rennes

Method 2 : block matching = similarity function

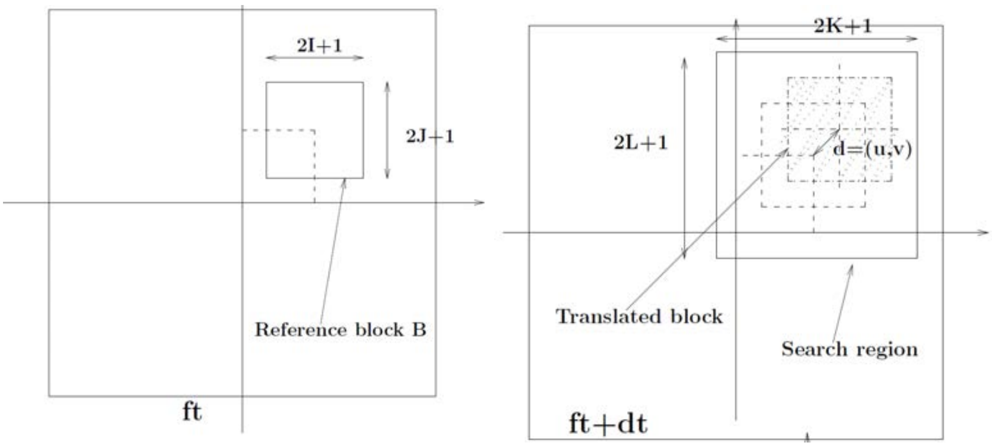


- L_1 norm = Sum of Absolute Differences (SAD)

$$C(i,j,u,v) = \sum_{p=-I}^{+I} \sum_{q=-J}^{+J} |f_{t+dt}(i+u+p, j+v+q) - f_t(i+p, j+q)|$$

$$C(u^*, v^*) = \min_{(u,v)\in\{-K,+K\}\times\{-L,+L\}} C(u,v)$$

Method 2 : block matching = similarity function

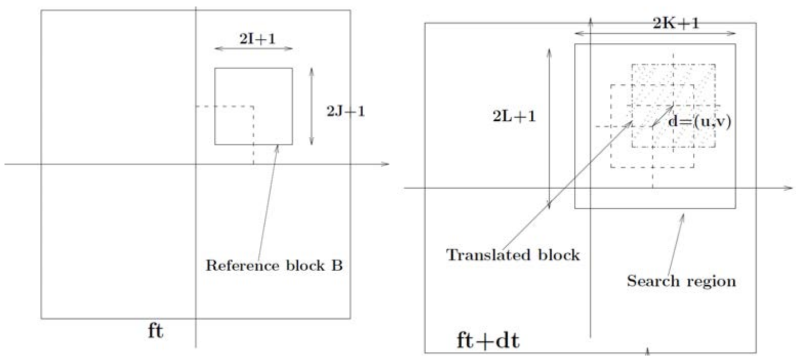


- L_2 norm = Sum of Squared Differences (SSD)

$$C(i,j,u,v) = \sum_{p=-I}^{+I} \sum_{q=-J}^{+J} (f_{t+dt}(i+u+p, j+v+q) - f_t(i+p, j+q))^2$$

$$C(u^*, v^*) = \min_{(u,v)\in\{-K,+K\}\times\{-L,+L\}} C(u,v)$$

Method 2 : block matching $\simeq$ correlation function

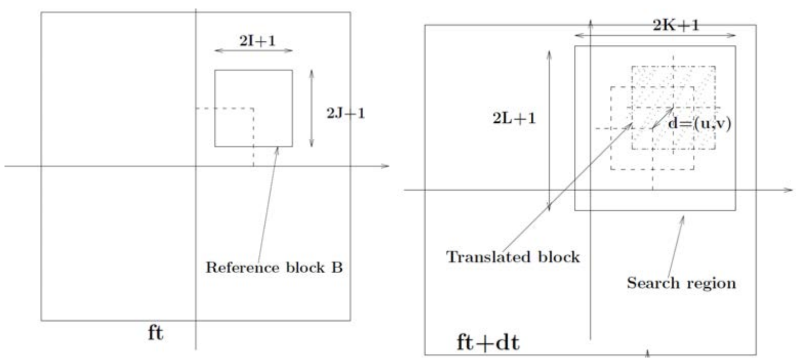


- Cross Correlation

$$C(i,j,u,v) = \sum_{p=-I}^{+I} \sum_{q=-J}^{+J} f_{t+dt}(i+u+p, j+v+q) \times f_t(i+p, j+q)$$

$$C(u^*, v^*) = \max_{(u,v)\in\{-K,+K\}\times\{-L,+L\}} C(u,v)$$

Method 2 : block matching $\simeq$ correlation function



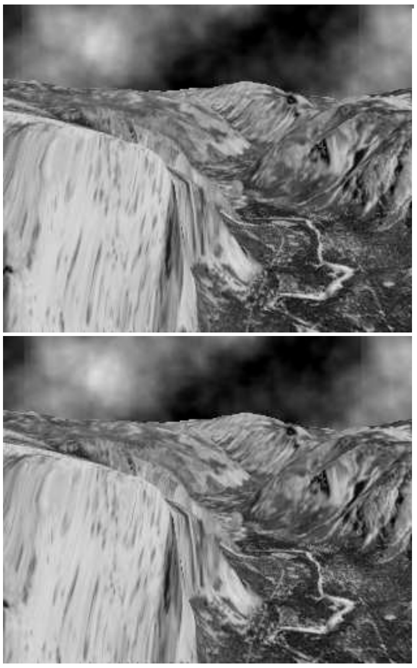
- Normalized Cross Correlation

$$C(i,j,u,v) = \frac{\sum_p \sum_q [f_{t+dt}(i+u+p, j+v+q) - \bar{f}_{t+dt}(i+u, j+v)][f_t(i+p, j+q) - \bar{f}_t(i,j)]}{\sqrt{\sigma^2_{f_{t+dt}}(i+u, j+v)}\sqrt{\sigma^2_{f_t}(i,j)}}$$

where

$$\sigma^2_f(x,y) = \frac{1}{(2I+1)(2J+1)} \sum_{p=-I}^{I} \sum_{q=-J}^{J} (f(x+p, y+q) - \bar{f}(x,y))^2$$

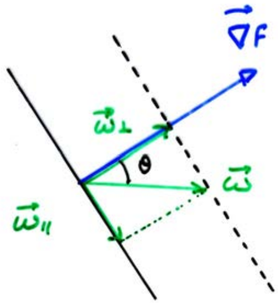
$$\bar{f}(x,y) = \frac{1}{(2I+1)(2J+1)} \sum_{p=-I}^{I} \sum_{q=-J}^{J} f(x+p, y+q)$$



C. Sun, Fast Optical Flow Using Cross Correlation and Shortest-Path Techniques, Digital Image Computing : Techniques and Applications. Perth, Australia, December 7-8, 1999, pp.143-148.

C. Sun, Fast Optical Flow Using Cross Correlation and Shortest-Path Techniques, Digital Image Computing : Techniques and Applications. Perth, Australia, December 7-8, 1999, pp.143-148.

Method 3 : first-order development of the DFD



- Small displacements : $d = (dx, dy)$, velocity vector : $\omega = (\frac{dx}{dt}, \frac{dy}{dt})$

$$\begin{aligned} DFD(p,d,t) &= f(p+d, t+dt) - f(p,t) \\ &= f(x+dx, y+dy, t+dt) - f(x,y,t) \end{aligned}$$

- First order development of DFD :

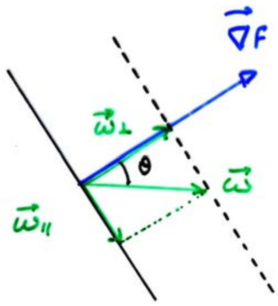
$$f(x+dx, y+dy, t+dt) = f(x,y,t) + dx\,\frac{\partial f}{\partial x} + dy\,\frac{\partial f}{\partial y} + dt\,\frac{\partial f}{\partial t} + h.o.t.$$

$$DFD(p,d,t) \simeq dx\,\frac{\partial f}{\partial x} + dy\,\frac{\partial f}{\partial y} + dt\,\frac{\partial f}{\partial t}$$

- Image flow constraint equation :

$$DFD(p,d,t) = 0 \;\Rightarrow\; \nabla f \cdot \omega = -\frac{\partial f}{\partial t}$$

Method 3 : first-order development of the DFD



- Image flow constraint equation :

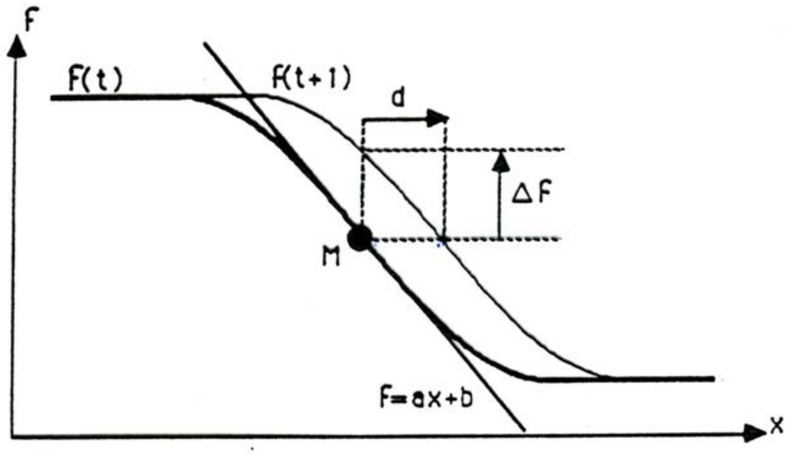
$$\nabla f(x,y,t) \cdot \omega(x,y,t) = -\frac{\partial f}{\partial t}(x,y,t)$$

- One equation, two unknowns : $\omega = (\frac{dx}{dt}, \frac{dy}{dt})$
- Valid for small displacements
- Only one component of velocity (or displacement) can be retrieved : $\omega_\perp$ ($\perp$ to the edge)

$$\nabla f \cdot \omega = \|\nabla f\|\,\|\omega\| \cos\theta = \|\nabla f\|\,\|\omega_\perp\| = -\frac{\partial f}{\partial t}$$

Method 3 : first-order development of the DFD

Interpretation



- First order development = local linearization of the intensity function
- One-dimensional case :

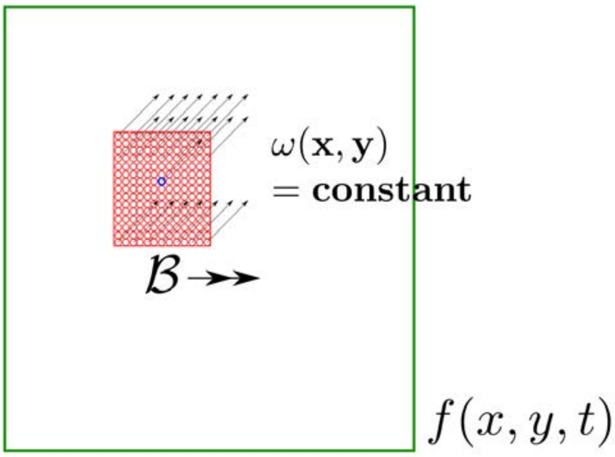
$$a \cdot d = -\Delta f$$

- Two-dimensional case :

$$\nabla f \cdot \omega = -\frac{\partial f}{\partial t}$$

Method 3 : first-order development of the DFD

Motion estimation using a local translation model

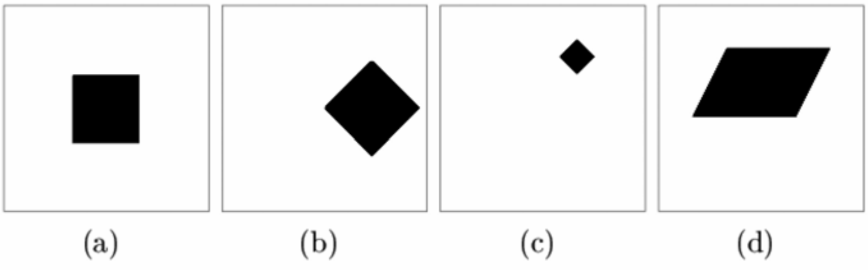


- Assumption : constant motion within block $\mathcal{B}$: $\omega(x, y) = (u, v) = constant$
- $N = Card(\mathcal{B})$ equations, two unknowns : (u, v) (Least-squares method)

$$\left\{ \begin{array}{rcl} \frac{\partial f}{\partial x}(x_1, y_1) \cdot u + \frac{\partial f}{\partial y}(x_1, y_1) \cdot v + \frac{\partial f}{\partial t}(x_1, y_1) & = & 0 \\ \frac{\partial f}{\partial x}(x_2, y_2) \cdot u + \frac{\partial f}{\partial y}(x_2, y_2) \cdot v + \frac{\partial f}{\partial t}(x_2, y_2) & = & 0 \\ \cdots & & \\ \frac{\partial f}{\partial x}(x_N, y_N) \cdot u + \frac{\partial f}{\partial y}(x_N, y_N) \cdot v + \frac{\partial f}{\partial t}(x_N, y_N) & = & 0 \end{array} \right.$$

- Estimated (u, v) is assigned to the (in blue) central point of sliding block $\mathcal{B}$.

Parametric motion models



(a) Original. (b) 2D rigid. (c) Similarity. (d) Affine.

Romero, F. Calderon, A Tutorial on Parametric Image Registration, I-Tech, Austria, 2007

- 2D translation

$$\omega(x,y) \;=\; \begin{cases} u(x,y) = t_x \\ v(x,y) = t_y \end{cases} \qquad \Theta = (t_x, t_y)$$

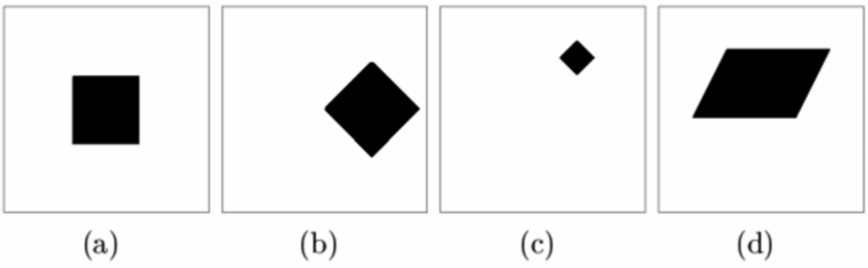
- 2D rigid motion

$$\omega(x,y) \;=\; \begin{cases} u(x,y) = (\cos\Phi - 1)\,x \;-\; \sin\Phi\, y \;+\; t_x \\ v(x,y) = \sin\Phi\, x \;+\; (\cos\Phi - 1)\, y \;+\; t_y \end{cases} \qquad \Theta = (\Phi, t_x, t_y)$$

- 2D similarity

$$\omega(x,y) \;=\; \begin{cases} u(x,y) = (s.\cos\Phi - 1)\,x \;-\; s.\sin\Phi\, y \;+\; s.t_x \\ v(x,y) = s.\sin\Phi\, x \;+\; (s.\cos\Phi - 1)\, y \;+\; s.t_y \end{cases} \qquad \Theta = (s, \Phi, t_x, t_y)$$

Parametric motion models



(a) Original. (b) 2D rigid. (c) Similarity. (d) Affine.

Romero, F. Calderon, A Tutorial on Parametric Image Registration, I-Tech, Austria, 2007

- Affine motion

$$\begin{cases} u(x,y) = a_1\, x \;+\; a_2\, y \;+\; t_x \\ v(x,y) = a_3\, x \;+\; a_4\, y \;+\; t_y \end{cases} \qquad \Theta = (a_1, a_2, a_3, a_4, t_x, t_y)$$

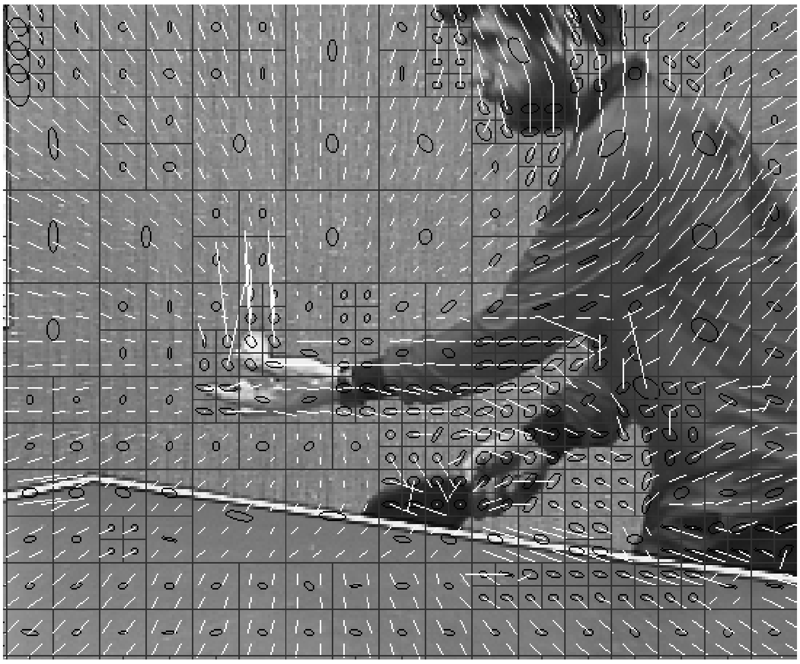
- Constrained quadratic model (perspective projection of the 3D rigid motion of a plane surface)

$$\begin{cases} u(x,y) = a_0 \;+\; a_1\, x \;+\; a_2\, y \;+\; a_6 x^2 \;+\; a_7\, xy \\ v(x,y) = a_3 \;+\; a_4\, x \;+\; a_5\, y \;+\; a_7 y^2 \;+\; a_6\, xy \end{cases} \qquad \Theta = (a_0, \cdots, a_7)$$

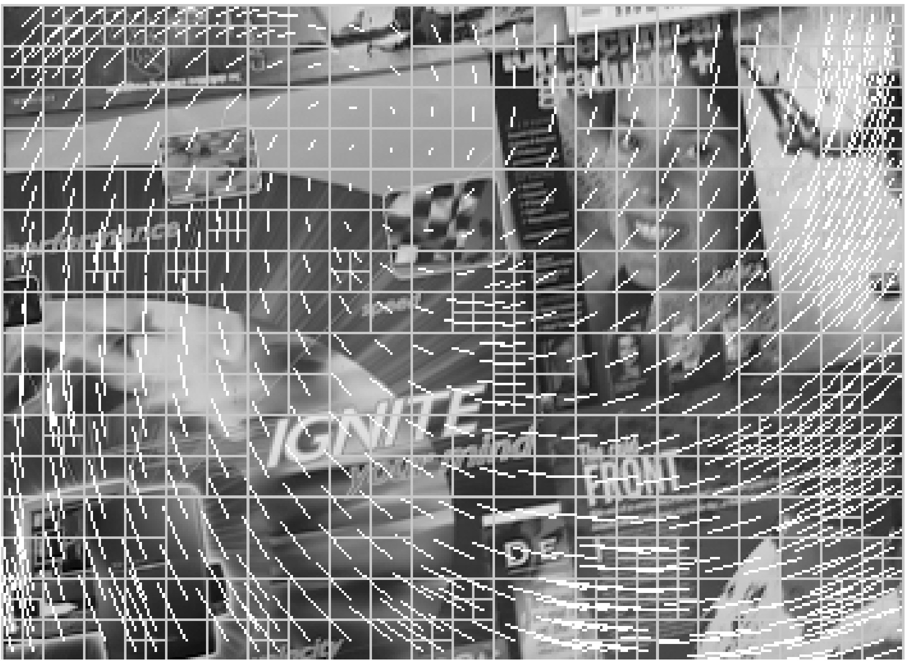
- Panoramic motion model (horizontal and vertical)

$$\begin{cases} u(x,y) = a_0 \;+\; a_6 x^2 \;+\; a_7\, xy \\ v(x,y) = a_3 \;+\; a_7 y^2 \;+\; a_6\, xy \end{cases} \qquad \Theta = (a_0, a_6, a_7)$$

- Splines deformation models, ...



S. Krüger, A. Calway, Motion Estimation and Tracking Using Multiresolution Affine Models, IEE Colloquium on Motion Analysis and Tracking, London, 1999



S. Krüger, A. Calway, Motion Estimation and Tracking Using Multiresolution Affine Models, IEE Colloquium on Motion Analysis and Tracking, London, 1999

Method 3 : first-order development of the DFD

Motion estimation using global quadratic regularization

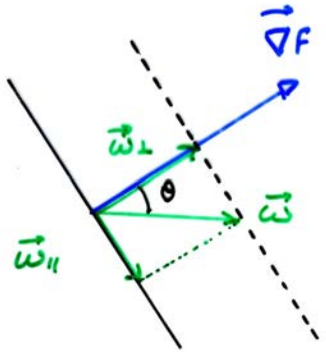
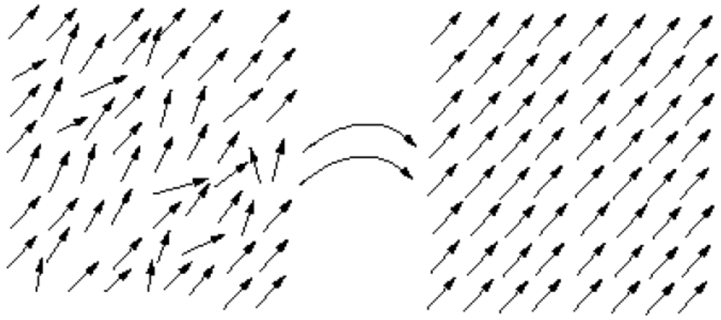


Image Flow Constraint

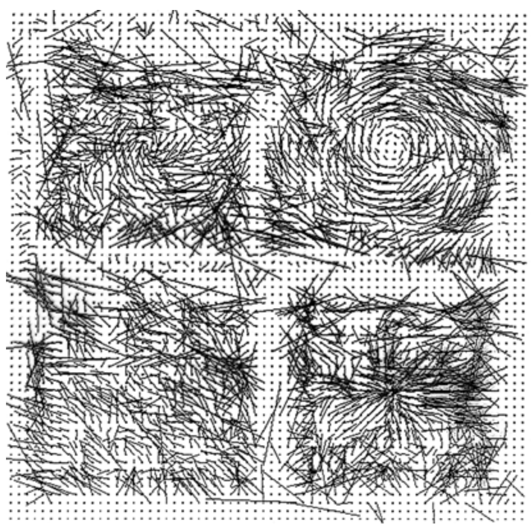


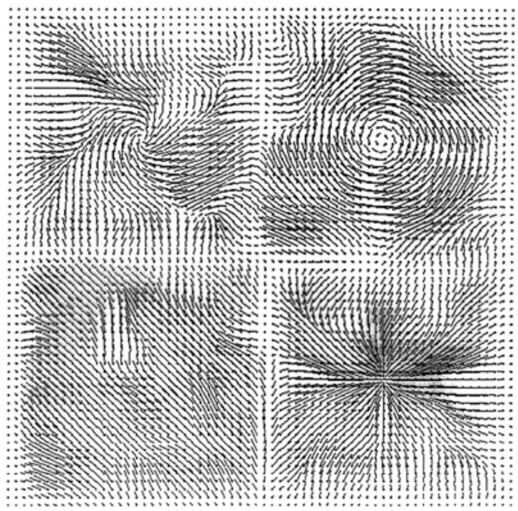
Regularization

- Estimation of a globally **smooth** motion field on image $\mathcal{I}$ through regularization

$$\min_{\{\omega(\boldsymbol{p},\boldsymbol{t}),\, \boldsymbol{p}\in\mathcal{I}\}} \sum_{\boldsymbol{p}\in\mathcal{I}} \left[\nabla f(p,t)\cdot\omega(p,t) + \frac{\partial f}{\partial t}\right]^2 + \alpha \sum_{\boldsymbol{p}\in\mathcal{I}} \|\nabla u(p,t)\|^2 + \|\nabla v(p,t)\|^2$$

  with $\omega(p,t) = (u(p,t), v(p,t))$, $\quad \alpha$ = regularization parameter
- Optimization : multiresolution relaxation techniques (iterative methods)


(a)


(b)

Motion estimation using global quadratic regularization (a) $\alpha \to 0$, $\quad$ (b) $\alpha \nearrow$

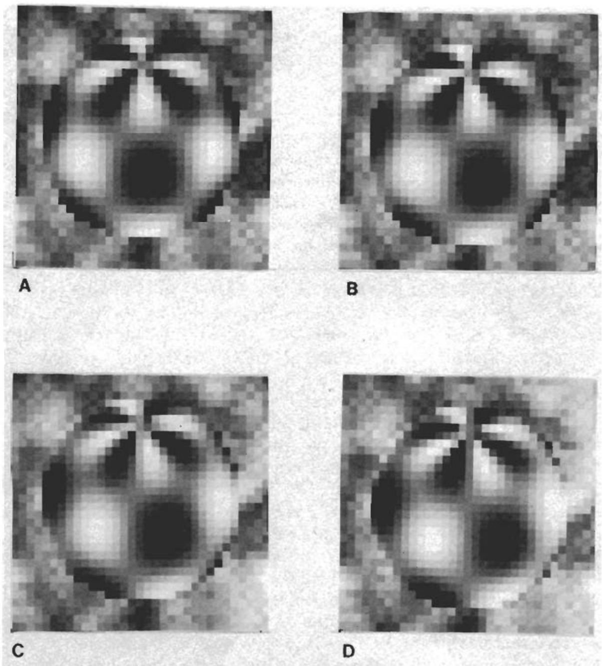


FIG. 5. Four frames out of a sequence of images of a sphere rotating about an axis inclined towards the viewer. The sphere is covered with a pattern which varies smoothly from place to place. The sphere is portrayed against a fixed, lightly textured background. Image sequences like these are processed by the optical flow algorithm.

B.K.P. Horn, B.G. Schunck, Determining Optical Flow, Artificial Intelligence, Vol. 17, pp. 185-203, 1981

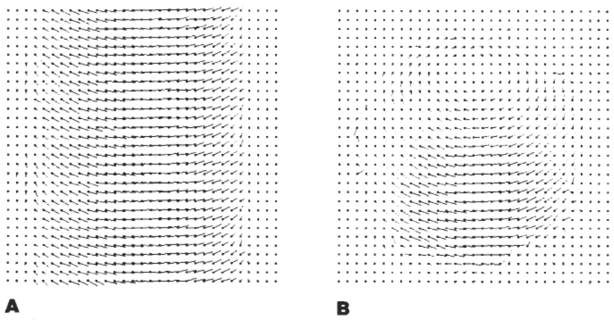


FIG. 10. Flow patterns computed for a cylinder rotating about its axis and for a rotating sphere. The axis of the cylinder is inclined 30 degrees towards the viewer and that of the sphere 45 degrees. Both are rotating at about 5 degrees per time step. The estimates shown are obtained after 32 time steps.

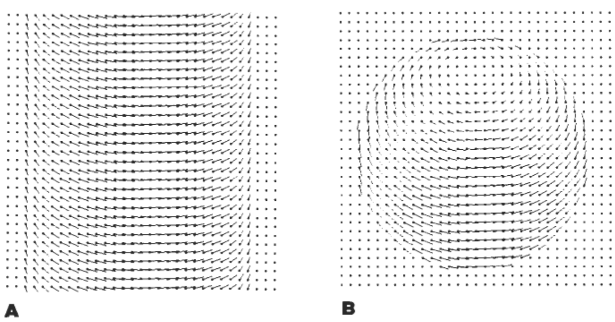
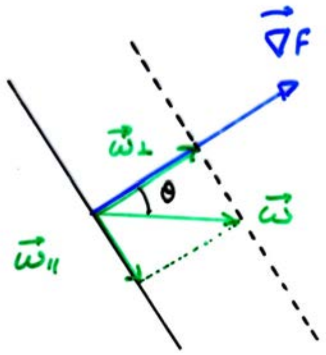


FIG. 11. Exact flow patterns for the cylinder and the sphere.

B.K.P. Horn, B.G. Schunck, Determining Optical Flow, Artificial Intelligence, Vol. 17, pp. 185-203, 1981

Method 3 : first-order development of the DFD

Motion estimation using global semi-quadratic regularization



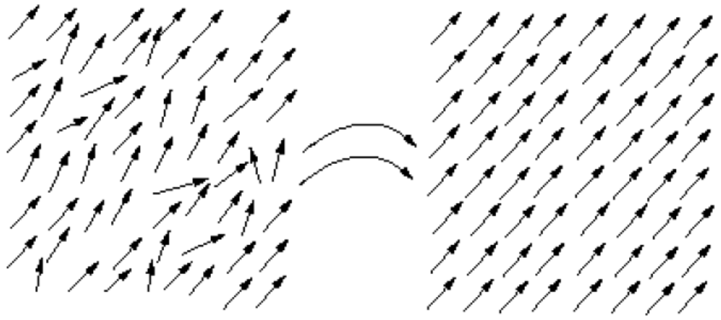


Image Flow Constraint **Regularization**

- Estimation of a globally **piece-wise smooth** motion field on image $\mathcal{I}$ through regularization

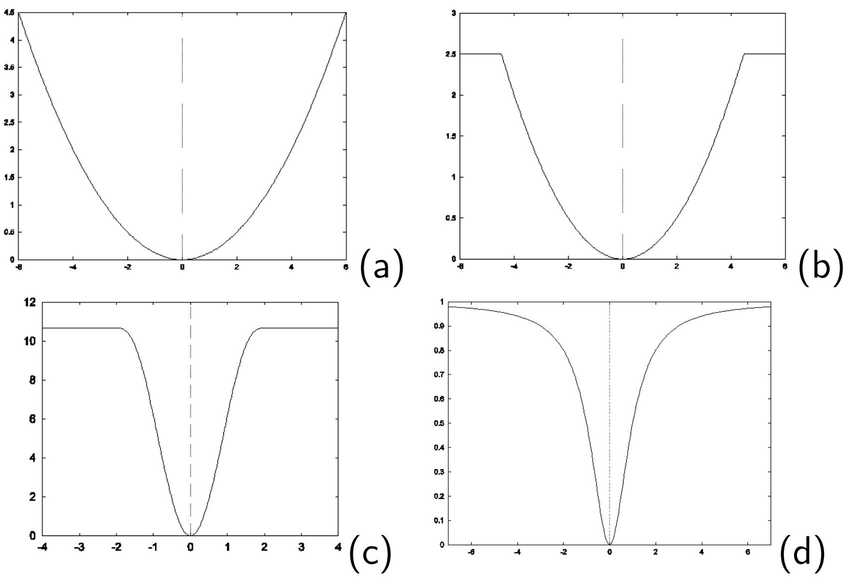
$$\min_{\{\omega(\boldsymbol{p},\boldsymbol{t}),\,\boldsymbol{p}\in\mathcal{I}\}} \sum_{\boldsymbol{p}\in\mathcal{I}} \rho_1\left[\nabla f(p,t)\cdot\omega(p,t)+\frac{\partial f}{\partial t}\right] + \alpha \sum_{\boldsymbol{p}\in\mathcal{I}} \rho_2\left(\|\omega(p,t)\|\right)$$

  ρ_1 and ρ_2 are robust (non quadratic) penalty functions that preserve discontinuities in the motion field

- Optimization : multiresolution relaxation techniques (iterative methods)

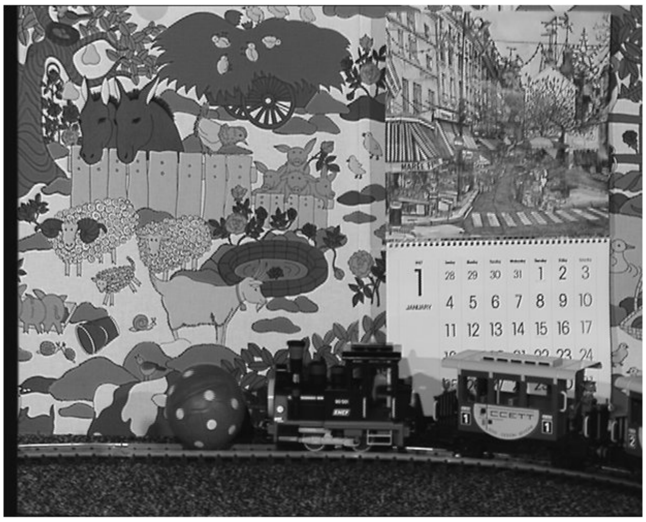
Method 3 : first-order development of the DFD

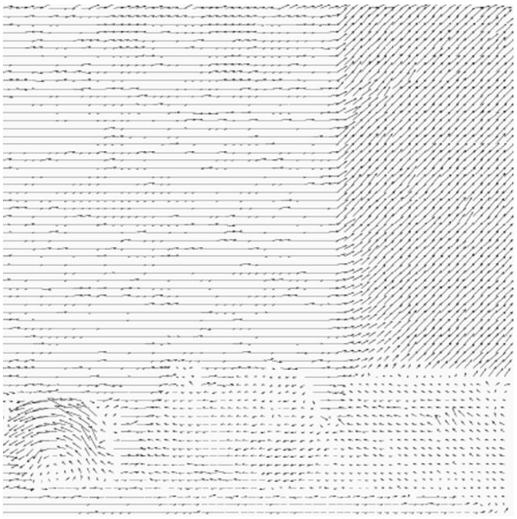
Robust penalty functions ρ



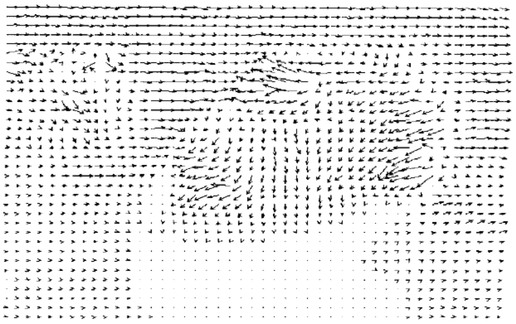
(a) Quadratic (non robust). (b) Truncated quadratic. (c) Tukey function. (d) Geman-McClure function.

$$\min_{\{\omega(\boldsymbol{p},\boldsymbol{t}),\,\boldsymbol{p}\in\mathcal{I}\}} \sum_{\boldsymbol{p}\in\mathcal{I}} \rho_1\left[\nabla f(p,t)\cdot\omega(p,t)+\frac{\partial f}{\partial t}\right] + \alpha \sum_{\boldsymbol{p}\in\mathcal{I}} \rho_2\left(\|\omega(p,t)\|\right)$$

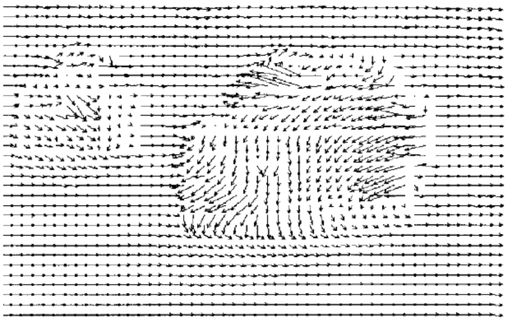




E. Memin, P. Perez, Dense Estimation and Object-Based Segmentation of the Optical Flow with Robust Techniques, IEEE Trans. Image Processing, May 1998



(a)



(b)

(a) Quadratic penalty function (smooth). (b) Non quadratic penalty function (preserves discontinuities)

E. Memin, P. Perez, Dense Estimation and Object-Based Segmentation of the Optical Flow with Robust Techniques, IEEE Trans. Image Processing, May 1998

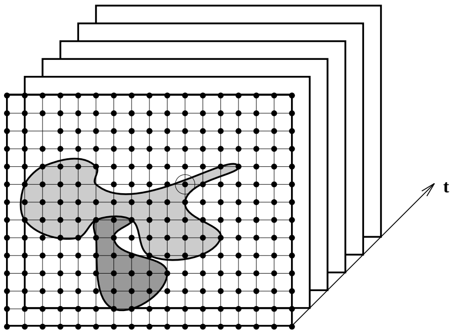
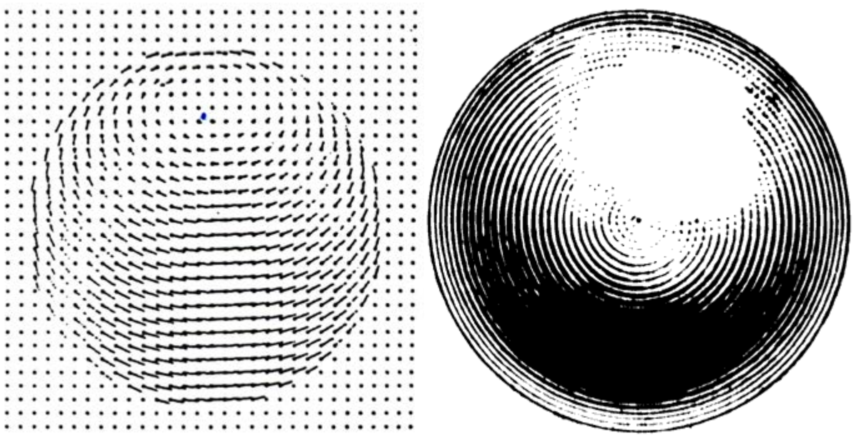


IMAGE SEQUENCE ANALYSIS

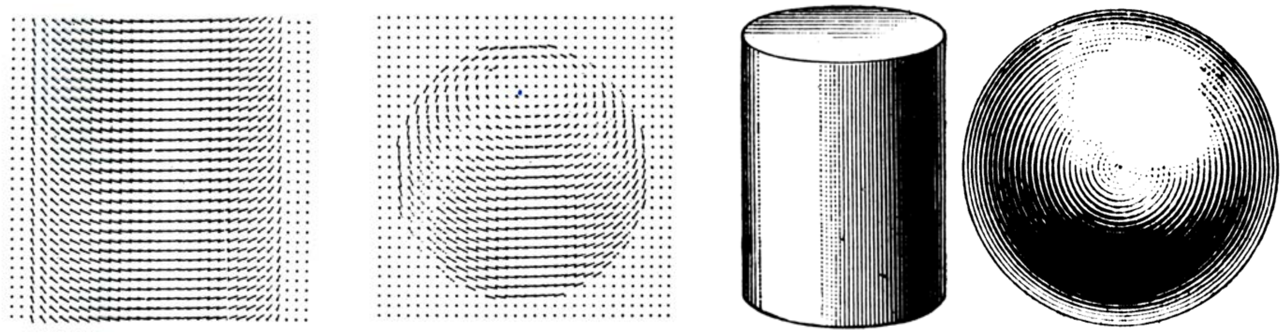
Part 3

IMAGE SEQUENCE ANALYSIS

3D structure and motion



3D structure and motion

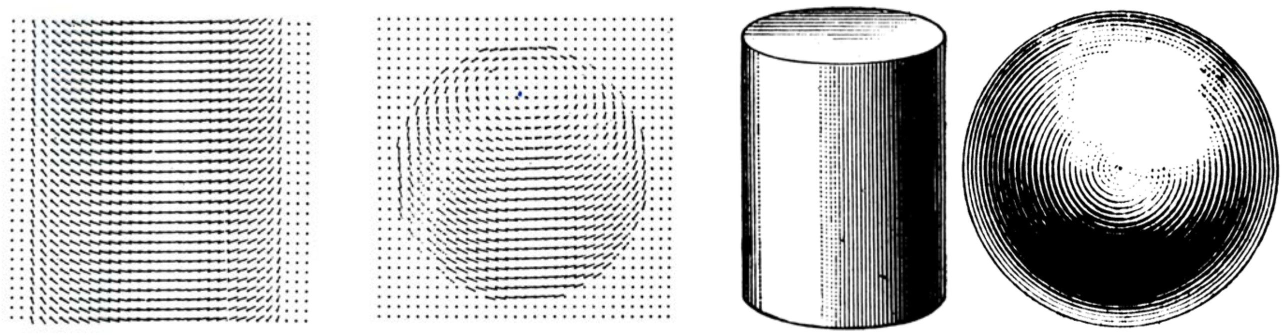


B.K.P. Horn, B.G. Schunck, Determining Optical Flow, Artificial Intelligence, Vol. 17, pp. 185-203, 1981

Clipart courtesy FCIT - http ://etc.usf.edu/clipart

- Goal : reconstruction of the 3D structure of a scene, reconstruction of 3D motion descriptors (3D rigid motion case).
- Assumption : 3D rigid motion (relative motion between scene and camera)
- Approaches :
  - from the 2D velocity field (in the image plane) ;
  - from the 2D displacement field (in the image plane)

3D structure and motion

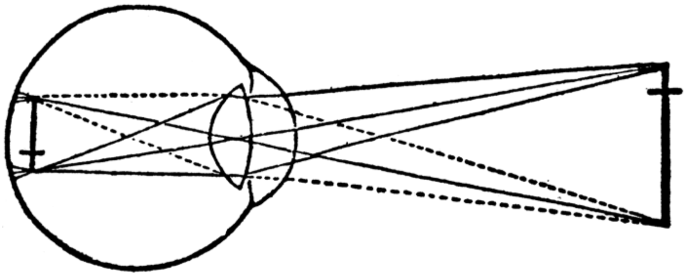


B.K.P. Horn, B.G. Schunck, Determining Optical Flow, Artificial Intelligence, Vol. 17, pp. 185-203, 1981

Clipart courtesy FCIT - http ://etc.usf.edu/clipart

- Several classes of problems :
  - stereo from motion, 3D shape from 2D motion : known rigid camera motion, static scene, determining 3D structure
  - 3D structure and motion (1 relative motion) : unknown rigid camera motion, static scene, determining 3D structure and motion
  - 3D structure and motion (N relative motions) : unknown rigid camera motion, rigid moving objects, determination 3D structure and multiple motions
  - calibrated / uncalibrated camera.

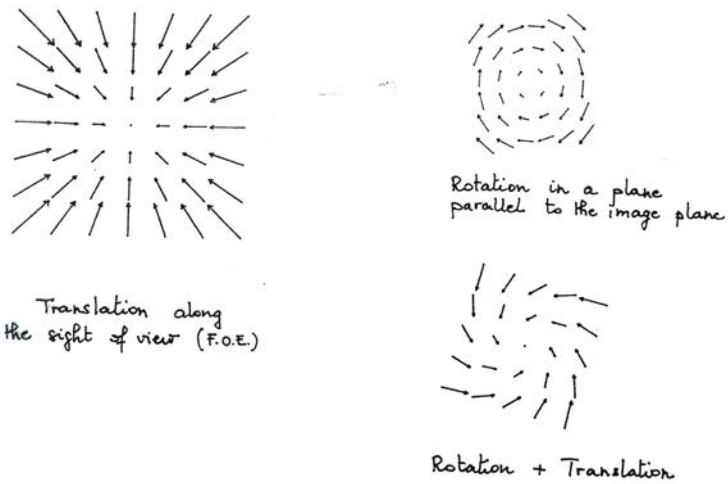
Elements of visual perception



Clipart courtesy FCIT - http ://etc.usf.edu/clipart

- The 2D velocity field on the retina provides information on the 3D motion and structure of a scene
- When objects are rigid, the interpretation of the relative motion between eye and scene is unique (same for the structure).
- Decomposition of 3D rigid motion :
  - the 3D information is only obtained from 3D translations
  - if the 3D motion is a translation, the 2D velocity field is divergent : all the vectors converge towards a single point (the Focus of Expansion).

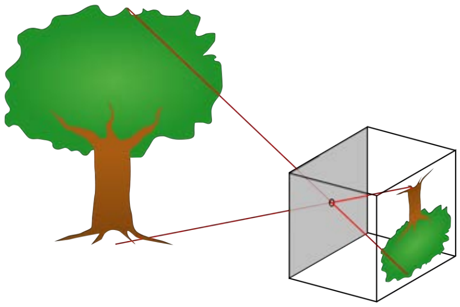
Elements of visual perception



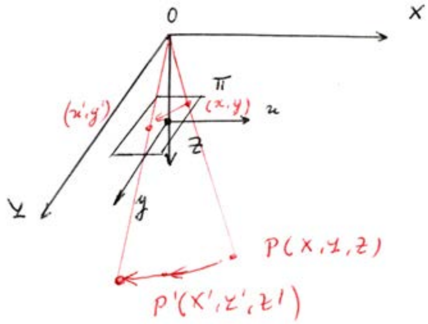
- Decomposition of 3D rigid motion :
  - the 3D information is only obtained from 3D translations
  - if the 3D motion is a translation, the 2D velocity field is divergent : all the vectors converge towards a single point (the Focus of Expansion).

3D structure and motion equations

Pinhole camera model



©Wikipedia



$P(X, Y, Z)$ projects on $p(x, y)$ according to :

- Pinhole camera model (optical center behind the image plane Π) :

$$x = f.\frac{X}{Z}$$
$$y = f.\frac{Y}{Z}$$

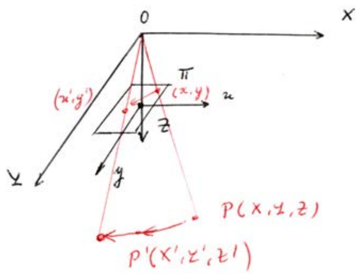
  where O is the optical center, f is the focal distance $= d(\Pi, O)$.

- Simplified pinhole camera model ($f = 1$) :

$$x = \frac{X}{Z}$$
$$y = \frac{Y}{Z}$$

3D structure and motion equations

Equations in position



$P(X, Y, Z)$ projects on $p(x, y)$. P moves to P' between t and $t + \Delta t$. The 3D rigid motion is characterized by :

- A 3×3 orthonormal rotation matrix : $(\mathbf{R}^{\mathrm{T}}\mathbf{R} = I, \det\mathbf{R} = 1)$ and a 3D translation vector :

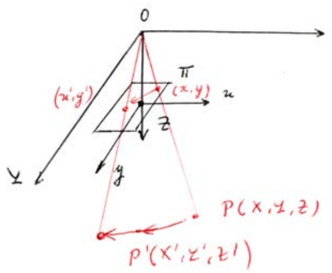
$$\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \qquad \mathbf{T} = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

- Rigid motion equation (in position) :

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \mathbf{R} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{T}$$

3D structure and motion equations

Equations in position



- Projection equations :

$$\begin{cases} x = \frac{X}{Z} \\ y = \frac{Y}{Z} \end{cases} \qquad \begin{cases} x' = \frac{X'}{Z'} \\ y' = \frac{Y'}{Z'} \end{cases}$$

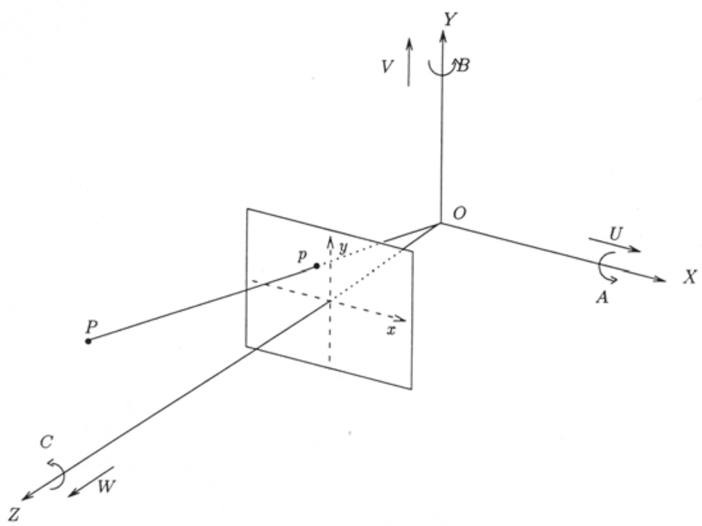
- 2D-3D relations :

$$\begin{cases} x' = \frac{(r_1 x + r_2 y + r_3)Z + T_X}{(r_7 x + r_8 y + r_9)Z + T_Z} \\ y' = \frac{(r_4 x + r_5 y + r_6)Z + T_Y}{(r_7 x + r_8 y + r_9)Z + T_Z} \end{cases}$$

  Unknowns : depth Z, motion parameters : r_i (only 3 unknowns ! 3 Euler angles), T_X, T_Y, T_Z
- In theory correspondences for 6 non-coplanar points are sufficient to solve the non-linear system.
- The translation vector can only be determined up to a scale factor (if Z, $\mathbf{T}$ is solution, αZ, $\alpha \mathbf{T}$ is also solution).

3D structure and motion equations

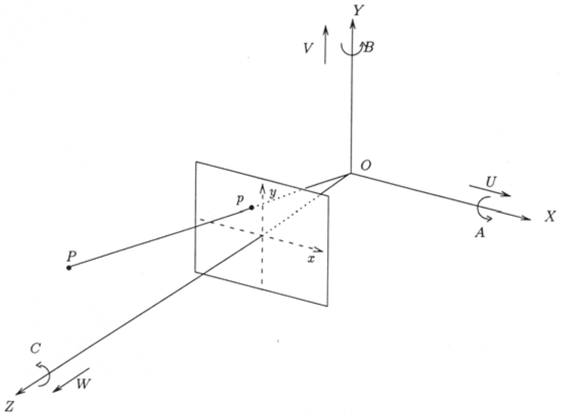
Equations in velocity



$$\mathbf{OP} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \mathbf{V} = \frac{d\mathbf{OP}}{dt} = \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} \qquad \mathbf{Op} = \begin{bmatrix} x = \frac{X}{Z} \\ y = \frac{Y}{Z} \\ f = 1 \end{bmatrix} \qquad \omega = \frac{d\mathbf{Op}}{dt} = \begin{bmatrix} u = \dot{x} \\ v = \dot{y} \\ 0 \end{bmatrix}$$

3D structure and motion equations

Equations in velocity



- Kinematics of rigid body (screw theory) :

$$\mathbf{V} = \frac{d\mathbf{OP}}{dt} = \Omega \times \mathbf{OP} + \mathbf{T}$$

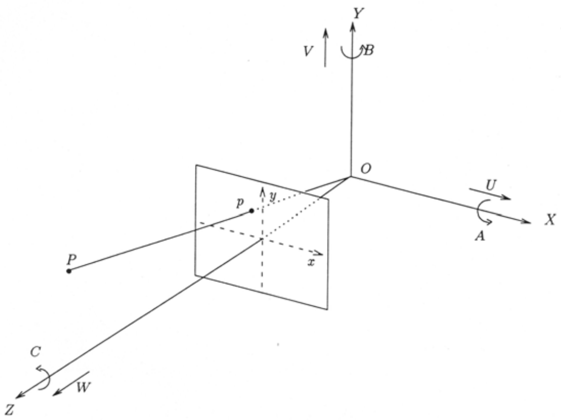
where :

$$\mathbf{T} = \begin{bmatrix} U \\ V \\ W \end{bmatrix} \qquad \Omega = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

are the translation velocity vector and the angular velocity vector (resp.).

3D structure and motion equations

Equations in velocity



- 2D-3D relations :

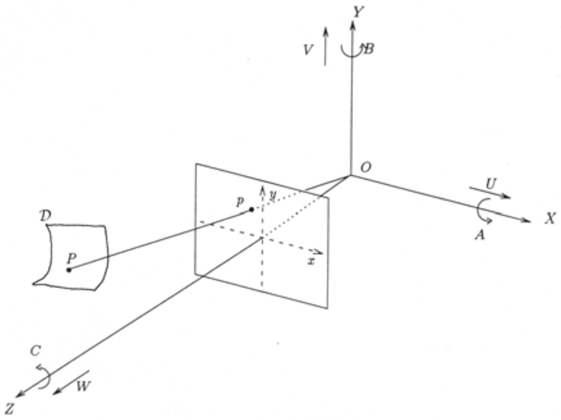
$$\omega = \begin{cases} u = \dot{x} = \frac{U - xW}{Z} - Axy + B(1+x^2) - Cy \\ v = \dot{y} = \frac{V - yW}{Z} - A(1+y^2) + Bxy + Cx \end{cases}$$

Unknowns : depth Z, U, V, W, A, B, C

- In theory the 2D velocities for 6 non-coplanar points are sufficient to solve the non-linear system.
- The translation velocity vector can only be determined up to a scale factor (if Z, $\mathbf{T}$ is solution, αZ, $\alpha\mathbf{T}$ is also solution).

3D structure and motion equations

Equations in velocity

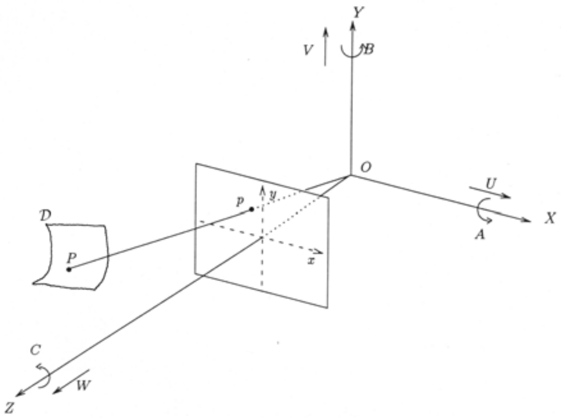


- Measurements : velocity field $\omega(x, y) \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$ for (X, Y, Z) in a region $\mathcal{D}$ of the scene ($\mathcal{D}$ projects onto D_{proj} in the image plane).
- Direct resolution

$$\min_{U,V,W,A,B,C,\{Z(x,y)\}} \quad \sum_{(x,y)\in D_{proj}} \left[\left(u(x,y) - \left(\frac{U - xW}{Z(x,y)} - Axy + B(1+x^2) - Cy \right) \right)^2 + \left(v(x,y) - \left(\frac{V - yW}{Z(x,y)} - A(1+y^2) + Bxy + Cx \right) \right)^2 \right]$$

3D structure and motion equations

Equations in velocity



- Local Taylor expansion of the surface $Z(X, Y)$ on region $\mathcal{D}$ in the scene.
  ▷ Planar (first order) approximation : $Z(X, Y) = \alpha X + \beta Y + Z_0$.
  ▷ Second order approximation : $Z(X, Y) = \alpha_1 X + \beta_1 Y + \alpha_2 X^2 + \beta_2 Y^2 + \gamma XY + Z_0$
- Example : planar surface. Unknowns : 3 (α, β, Z_0) + 6 (U, V, W, A, B, C)

$$Z = \alpha X + \beta Y + Z_0 \quad \Longrightarrow \quad \frac{1}{Z} = \frac{1}{Z_0}(1 - \alpha x - \beta y)$$

$$\begin{cases} u(x,y) = (U - xW)(\frac{1}{Z_0}(1 - \alpha x - \beta y)) - Axy + B(1+x^2) - Cy \\ v(x,y) = (V - yW)(\frac{1}{Z_0}(1 - \alpha x - \beta y)) - A(1+y^2) + Bxy + Cx \end{cases}$$

which is a constrained quadratic model for the velocity field !

Télécom Physique Strasbourg

Module 36032

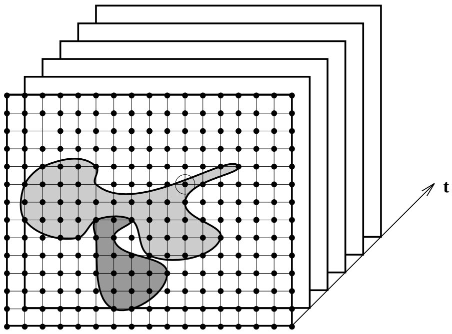
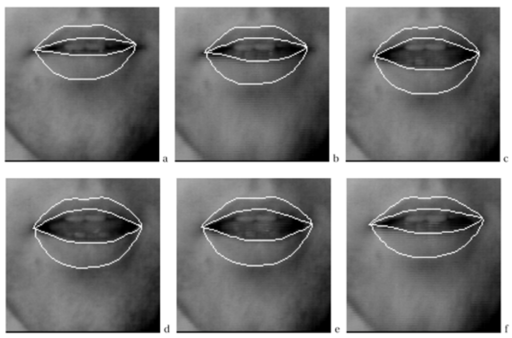


IMAGE SEQUENCE ANALYSIS

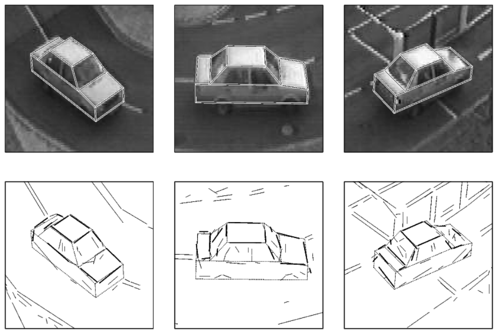
Part 4

IMAGE SEQUENCE ANALYSIS

Tracking



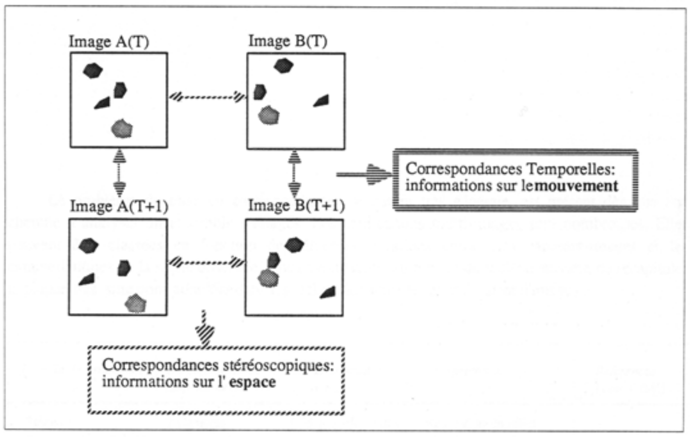
Tracking



D. Koller, K. Daniilidis, and H.-H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, IJCV, vol. 10, pp. 257-281, July 1993.

- Goal : tracking of a segmented structure over a long image sequence, trajectography.
- Approaches :
  - short term tracking : correspondence of image features, block matching ;
  - long term tracking with prediction (linear system) : recursive temporal filtering (Kalman filter) [Crowley, Deriche 1985] ;
  - long term tracking with prediction (non-linear system) : particle filtering [Isard and Blake 1996]

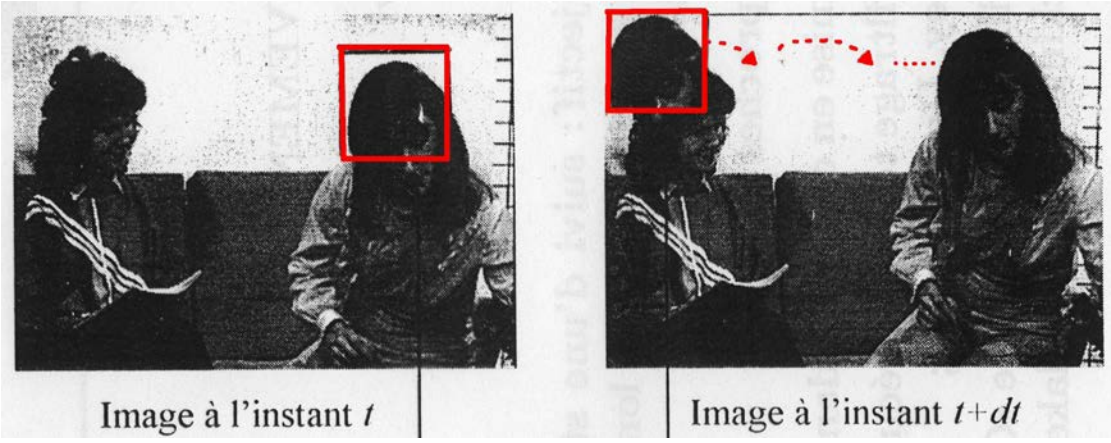
Short term tracking : correspondence



L. Pocztar, Thèse Univ. Paris 6, 1987

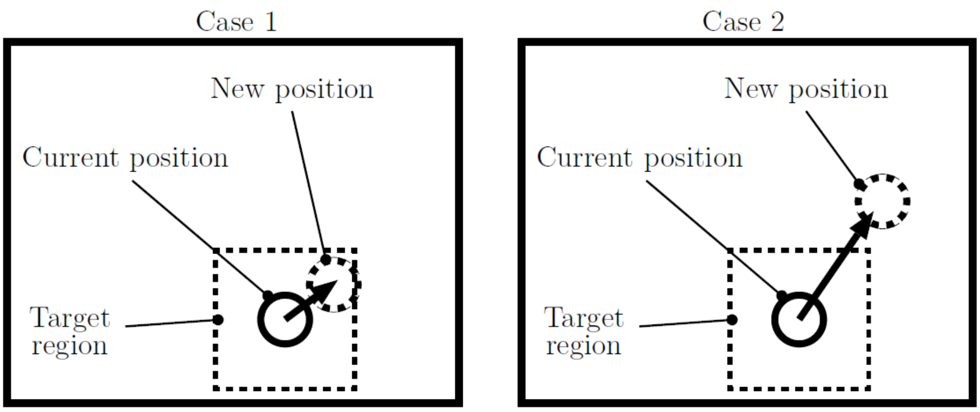
- Matching of discrete image features from time t to $t+1$.
- Approaches :
  - matching of points of interest (Moravec, SIFT), edges, regions, graphs
  - template or block matching, correlation ;
- Limitations : large movements (need exhaustive search) , wrong correspondence (periodic or repetitive patterns, ambiguous correspondences).

Short term tracking : correspondence



IRISA/INRIA Rennes

Short term tracking : correspondence

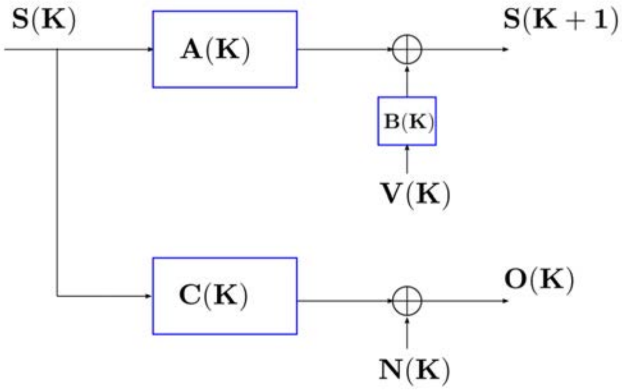


Case 1 : Tracking the object without position prediction might be successful,

Case 2 : Tracking without position prediction will fail

from N. Funk, University of Alberta, Project for CMPUT 652, December 7, 2003

Long term tracking : the Kalman filter

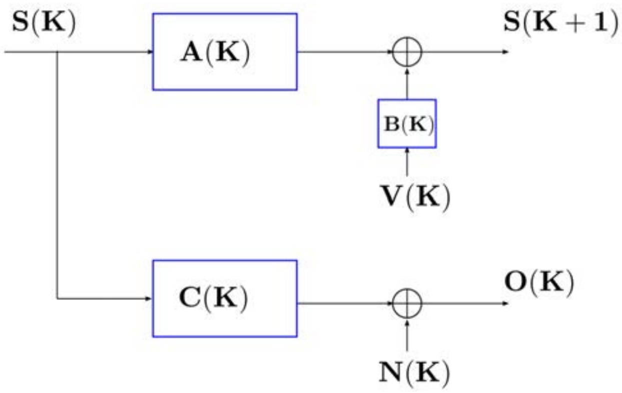


State-space representation of a linear system

- Non stationary linear system, gaussian noise
- The filter incorporates a kinematic model
- The kinematic model enables prediction of movement (from current position $\longrightarrow$ new position).
- The filter takes into account errors in the kinematic model and measurement noise.
- Efficient, low-cpu cost, recursive implementation (real time)

The Kalman filter

State-space representation of a linear system



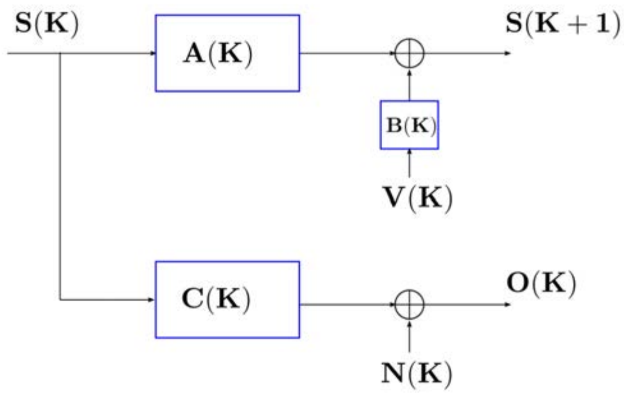
$S(k)$: state vector at time k
$O(k)$: measurement or observation vector at time k

$$\left\{ \begin{array}{lcl} S(k+1) & = & A(k)S(k) + B(k)V(k) \\ O(k) & = & C(k)S(k) + N(k) \end{array} \right. \qquad \begin{array}{l} \text{State transition equation} \\ \text{Measurement equation} \end{array}$$

$V(k)$: state transition noise
$N(k)$: observation noise
$A(k)$: state transition matrix, $B(k)$: state transition noise matrix , $C(k)$: observation matrix

The Kalman filter

State-space representation of a linear system

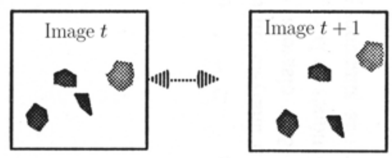


Hypotheses :

- $V(k)$ is a vector zero mean Gaussian white noise (not correlated with state $S(k)$)
  $\mathbb{E}[V(k)] = 0$
  $\mathbb{E}[V(k)V(j)^T] = \delta_{kj}\Sigma_V$ state transition noise covariance matrix
- $N(k)$ vector zero mean Gaussian white noise (not correlated with noise $V(k)$)
  $\mathbb{E}[N(k)] = 0$
  $\mathbb{E}[N(k)N(j)^T] = \delta_{kj}\Sigma_N$ observation noise covariance matrix

The Kalman filter

State-space representation of a linear system : example (constant velocity model)



$S(t)$ = position, speed, acceleration at time t : $S(t) = \begin{bmatrix} x(t) \\ y(t) \\ \dot{x}(t) \\ \dot{y}(t) \end{bmatrix}$ or $S(t) = \begin{bmatrix} x(t) \\ y(t) \\ \dot{x}(t) \\ \dot{y}(t) \\ \ddot{x}(t) \\ \ddot{y}(t) \end{bmatrix}$ ...

$O(t)$: measurement of position $= \begin{bmatrix} x_{mes}(t) \\ y_{mes}(t) \end{bmatrix}$ at time t

Kinematic model example : constant velocity model

$$\begin{cases} S(t+1) = A(t)S(t) + B(t)V(t) & \text{kinematic model} \\ O(t) = C(t)S(t) + N(t) & \text{measurement equation} \end{cases}$$

with : $S(t) = \begin{bmatrix} x(t) \\ y(t) \\ \dot{x}(t) \\ \dot{y}(t) \end{bmatrix}$, $A(t) = A = \begin{bmatrix} I_2 & \Delta t\, I_2 \\ O_2 & I_2 \end{bmatrix}$, $B(t) = B = I_4$, $C(t) = [I_2, O_2]$

The Kalman filter

What is the Kalman filter ?

The Kalman filter is a *recursive linear estimator* E_l of $S(k)$ given the observations O.

At time k the filter computes 2 estimates of state $S(k)$:

- Prediction from $k-1 \rightarrow k$:
  $\hat{S}(k|k-1) = E_l[S(k) \mid O(k-1), O(k-2), ..., O(0)]$
- Update of state $S(k)$, after observation of $O(k)$:
  $\hat{S}(k|k) = E_l[S(k) \mid O(k), O(k-1), ..., O(0)]$

E_l is the optimal linear minimum mean square error (MMSE) estimator, i.e. it minimizes :
$\mathbb{E}[(S(k) - \hat{S}(k|k-1))^2]$
and :
$\mathbb{E}[(S(k) - \hat{S}(k|k))^2]$

The estimation is *recursive* : only the estimated state $S(k-1|k-1)$ from the previous time step $k-1$ and the current observation $O(k)$ are needed to compute the estimate for the current state $S(k|k)$.

The Kalman filter

What is the Kalman filter ?

Recursive Kalman filtering produces two kind of errors :

- Prediction error :
  $\tilde{S}(k|k-1) = S(k) - \hat{S}(k|k-1)$
  $\tilde{S}$ is zero mean with covariance matrix :
  $\Sigma_{\tilde{S}}(k|k-1) = \mathbb{E}[\tilde{S}(k|k-1)\tilde{S}(k|k-1)^T]$

- Estimation (update or filtering) error :
  $\tilde{S}(k|k) = S(k) - \hat{S}(k|k)$
  with covariance matrix : $\Sigma_{\tilde{S}}(k|k) = \mathbb{E}[\tilde{S}(k|k)\tilde{S}(k|k)^T]$

The Kalman filter

Kalman filter recursive equations

3-a Prediction step

$\hat{S}(k|k-1) = E_l[S(k) \mid O(k-1), O(k-2), \ldots, O(0)]$
$\hat{S}(k|k-1) = A(k-1)\,\hat{S}(k-1|k-1)$

Prediction error :
$\tilde{S}(k|k-1) = S(k) - \hat{S}(k|k-1)$
$\Sigma_{\tilde{\mathbf{S}}}(k|k-1) = A(k-1)\Sigma_{\tilde{\mathbf{S}}}(k-1|k-1)A(k-1)^{\mathbf{T}} + B(k-1)\Sigma_{\mathbf{V}}B(k-1)^{\mathbf{T}}$

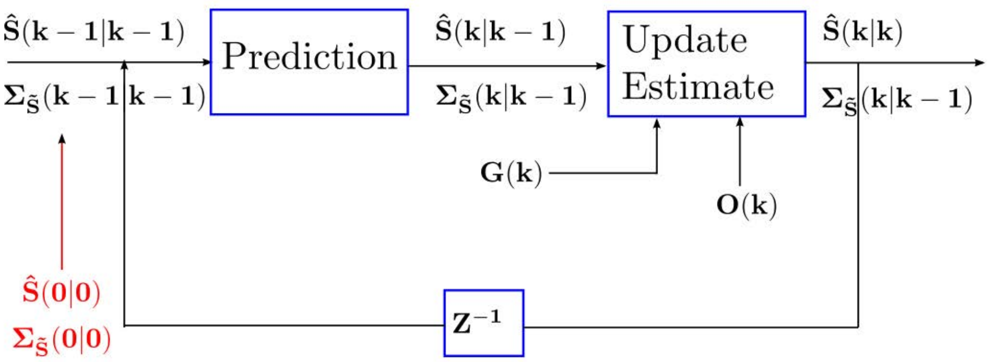
3-b Update (filtering) step : $O(k)$ is observed !

$\hat{S}(k|k) = E_l[S(k) \mid O(k), O(k-1), \ldots, O(0)]$
$\hat{S}(k|k) = \hat{S}(k|k-1) + G(k)[O(k) - C(k)\hat{S}(k|k-1)]$
where $G(k)$ is the Kalman filter gain :
$G(k) = \Sigma_{\tilde{\mathbf{S}}}(k|k-1)C(k)^{\mathbf{T}}[C(k)\Sigma_{\tilde{\mathbf{S}}}(k|k-1)C(k)^{\mathbf{T}} + \Sigma_{\mathbf{N}}]^{-\mathbf{1}}$

Estimation error :
$\tilde{S}(k|k) = S(k) - \hat{S}(k|k)$
$\Sigma_{\tilde{\mathbf{S}}}(k|k) = [I - G(k)C(k)]\Sigma_{\tilde{\mathbf{S}}}(k|k-1)$

The Kalman filter

Kalman filter recursive equations

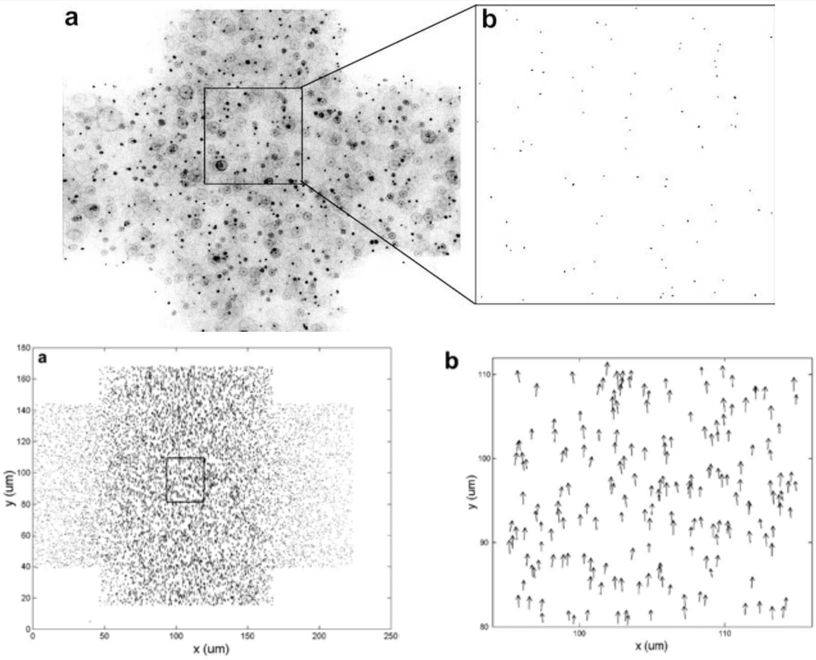


$$\hat{S}(k|k-1) = A(k-1)\,\hat{S}(k-1|k-1)$$
$$\Sigma_{\tilde{\mathbf{S}}}(k|k-1) = A(k-1)\Sigma_{\tilde{\mathbf{S}}}(k-1|k-1)A(k-1)^{\mathbf{T}} + B(k-1)\Sigma_{\mathbf{V}}B(k-1)^{\mathbf{T}}$$

$$G(k) = \Sigma_{\tilde{\mathbf{S}}}(k|k-1)C(k)^{\mathbf{T}}[C(k)\Sigma_{\tilde{\mathbf{S}}}(k|k-1)C(k)^{\mathbf{T}} + \Sigma_{\mathbf{N}}]^{-\mathbf{1}}$$
$$\hat{S}(k|k) = \hat{S}(k|k-1) + G(k)[O(k) - C(k)\hat{S}(k|k-1)]$$
$$\Sigma_{\tilde{\mathbf{S}}}(k|k) = [I - G(k)C(k)]\Sigma_{\tilde{\mathbf{S}}}(k|k-1)$$

Tracking using the Kalman filter

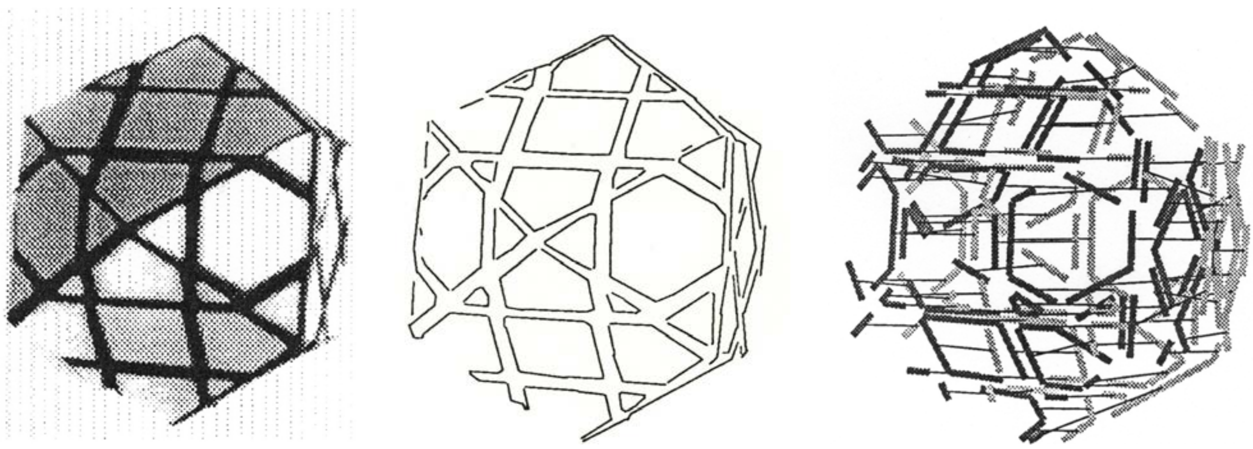
Particle tracking (particle velocimetry)



Particle tracking in fluid (particles : 500-nm diameter polystyrene spheres), algorithm : cross-correlation and Kalman filter prediction **S. Devasenathipathy, J.G. Santiago, S.T. Wereley, C.D. Meinhart, K. Takehara, Particle imaging techniques for microfabricated fluidic systems, Experiments in Fluids, Vol. 34 (2003) pp. 504-514**

Tracking using the Kalman filter

Edge lines tracking

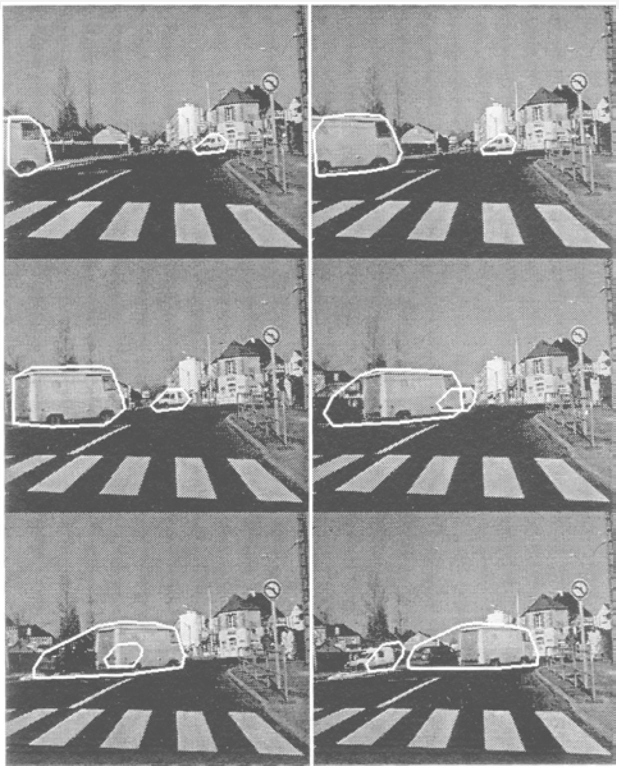


Tracking between image 5 (138 segments) and image 15 (161 segments). 86 segments have been tracked

J.L. Crowley, P. Stelmaszyk, C. Discours, Measuring Image Flow by Tracking Edge Lines, IEEE ICCV, 1988

Tracking using the Kalman filter

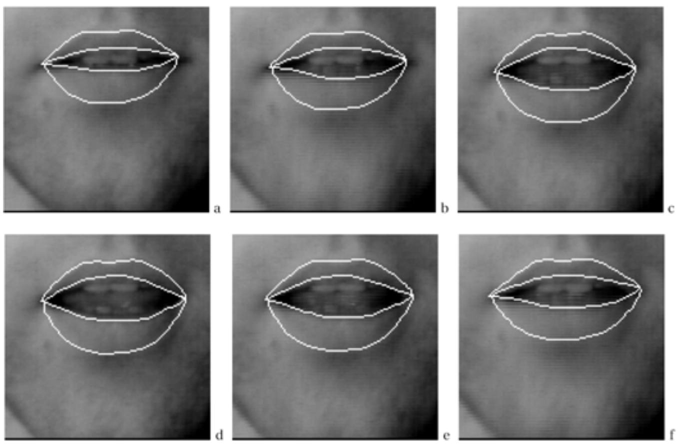
Region tracking



F. Meyer, Thèse Univ. Rennes I, 1993

Tracking using the Kalman filter

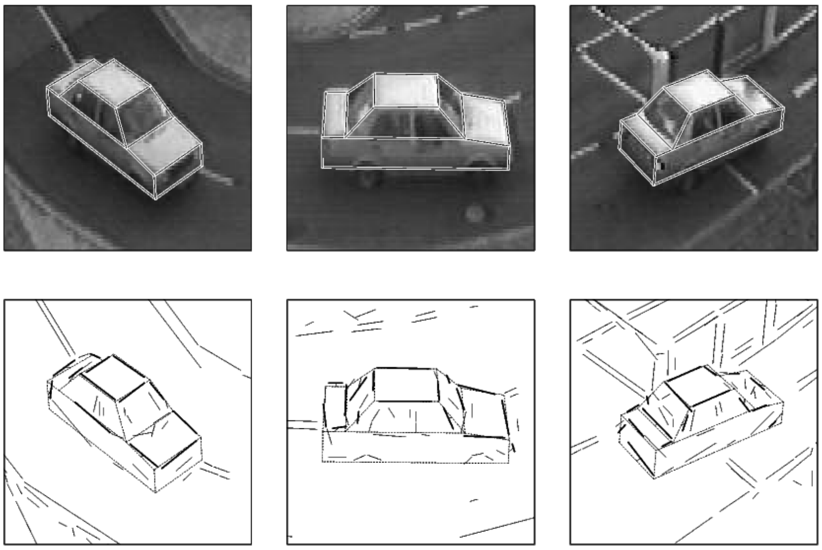
Deformable model tracking



C. Kervrann, F. Heitz, A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes,
CVGIP : Graphical Models and Image Processing, pp. 173-195, Vol. 60, Num. 3, 1998

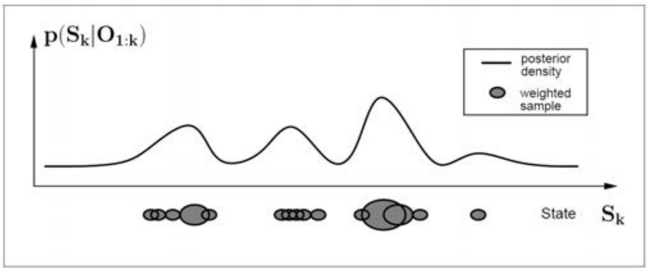
Tracking using the Kalman filter

3D model tracking



D. Koller, K. Daniilidis, and H.-H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, IJCV, vol. 10, pp. 257-281, July 1993.

Long term tracking : the particle filter



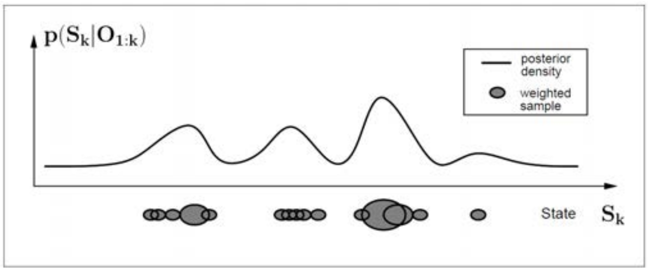
M. Isard, A. Blake, Conditional density propagation for visual tracking, Int. J. Computer Vision, Vol. 29, No 1, pp. 5-28, 1998

- Also known as : sequential Monte-Carlo method. Probabilistic formulation, stochastic sampling.
- In the most general case : non linear non stationary model, non gaussian distributions
- The Kalman filter may be considered as a particular case (linear system, gaussian noise)
- Sequential, recursive estimation of state
- Enables non-linear dynamics, multiple trajectories hypotheses, non gaussian noise...
- Far more time-consuming than Kalman filter

References :
- M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, IEEE Trans. on Signal Processing, Vol. 50, No. 2, 2002.
- M. Isard, A. Blake, Conditional density propagation for visual tracking, Int. J. Computer Vision, Vol. 29, No 1, pp. 5-28, 1998.
- A. Doucet and Adam M. Johansen, A tutorial on particle filtering and smoothing : fifteen years later ; in The Oxford Handbook of Nonlinear Filtering, OUP Oxford, 2011.

The particle filter

Non-linear non-Gaussian state space representation



$$\begin{cases} S_k = f_k(S_{k-1}, V_k) & \longleftrightarrow & p(S_k|S_{k-1}) \\ O_k = h_k(S_k, N_k) & \longleftrightarrow & p(O_k|S_k) \end{cases}$$

State transition equation
Measurement equation

S_k : state vector at time k (only depends on previous state S_{k-1})
O_k : measurement or observation vector at time k
V_k : state transition noise (white, non necessarily gaussian)
N_k : observation noise (white, non necessarily gaussian)
f_k, h_k : transition and measurement functions (eventually non-linear)

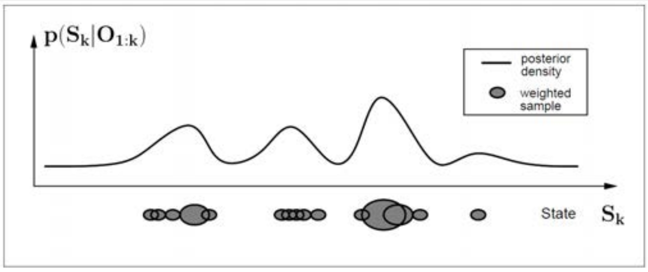
Estimation of state S_k
The particle filter is based on Monte-Carlo simulation, i.e. generation of samples according to the a posteriori distribution :

$$p(S_k|O_1, O_2, ..., O_k) = p(S_k|O_{1:k})$$

from which estimates of state S_k at time k may be derived. Standard estimators are the mode or the mean of this distribution.

The particle filter

Recursive estimation



Known :
$$\begin{cases} p(S_k|S_{k-1}) \text{ and } p(S_0) & \text{state transition probability} \\ p(O_k|S_k) & \text{measurement likelihoods} \end{cases}$$

- The particle filter is based on Monte-Carlo simulation, i.e. generation of samples according to the a posteriori distribution : $p(S_k|O_1, O_2, ..., O_k) = p(S_k|O_{1:k})$
- The recursive particle filter is based on a prediction and update step.
- Assume that $p(S_{k-1}|O_{1:k-1})$ has been calculated :

Prediction step

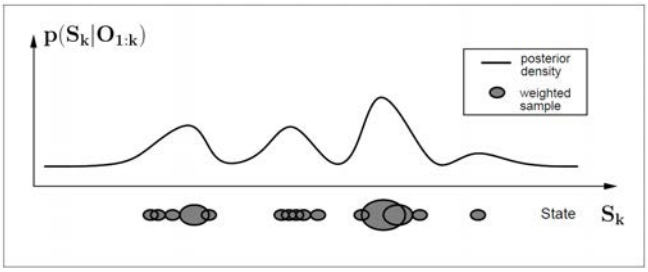
$$p(S_k|O_{1:k-1}) = \int p(S_k|S_{k-1}\, p(S_{k-1}|O_{1:k-1})\, dS_{k-1}$$

Update step
At time step k a measurement O_k becomes available.

$$p(S_k|O_{1:k}) = \frac{p(O_k|S_k)\, p(S_k|O_{1:k-1})}{\int p(O_k|S_k)\, p(S_k|O_{1:k-1})\, dS_k}$$

The particle filter

Recursive estimation



Prediction step

$$p(S_k|O_{1:k-1}) \;=\; \int p(S_k|S_{k-1}\; p(S_{k-1}|O_{1:k-1})\, dS_{k-1}$$

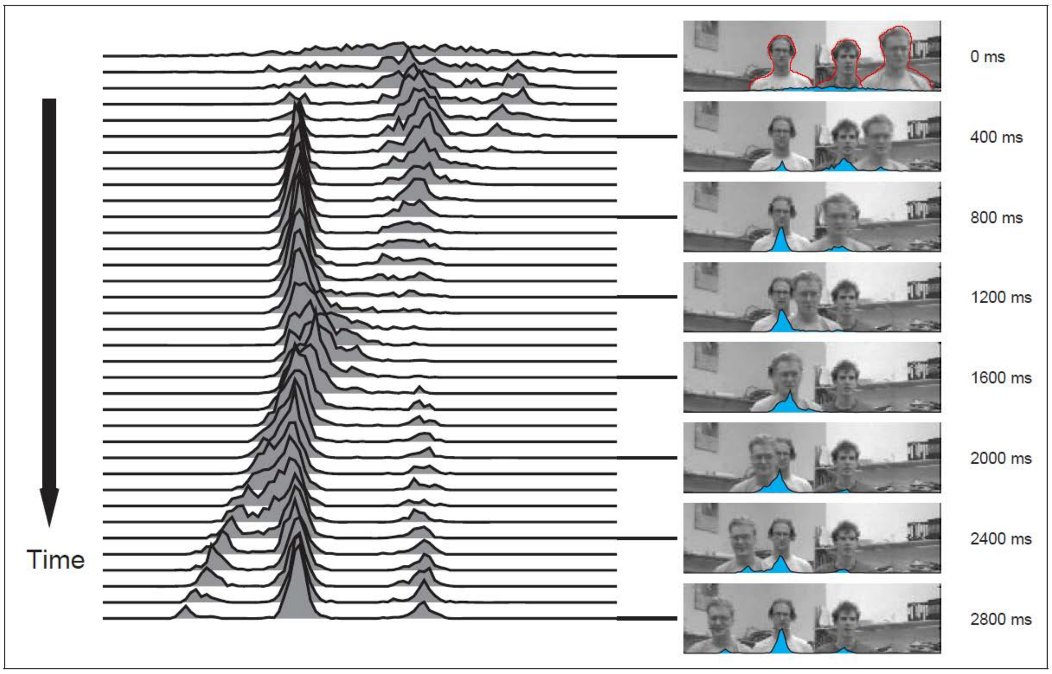
Update step

$$p(S_k|O_{1:k}) \;=\; \frac{p(O_k|S_k)\, p(S_k|O_{1:k-1})}{\int p(O_k|S_k)\, p(S_k|O_{1:k-1})\, dS_k}$$

"The recursive calculation of the posterior density $p(S_k|O_{1:k})$ cannot in general be determined analytically. This calculation is performed using sequential importance sampling (SIS) algorithms (Monte Carlo (MC) methods). The key idea is to represent the required posterior density function by a set of random samples with associated weights and to compute estimates based on these samples and weights. As the number of samples becomes very large, this MC characterization becomes an equivalent representation to the usual functional description of the posterior pdf, and the SIS filter approaches the optimal Bayesian estimate." (M.S. Arulampalam et al.)

Tracking using the particle filter

Template tracking



M. Isard, A. Blake, Conditional density propagation for visual tracking, Int. J. Computer Vision, Vol. 29, No 1, pp. 5-28, 1998.

Tracking using the particle filter

Template tracking

M. Isard, A. Blake, Conditional density propagation for visual tracking, Int. J. Computer Vision, Vol. 29, No 1, pp. 5-28, 1998.

Tracking using the particle filter

Color-based tracking

P. Pérez, J. Vermaak, A. Blake, Data Fusion for Visual Tracking with Particles, Proc. of the IEEE, Vol. 92, No. 3, 2004